

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: December 10, 2005, 23:23:15 ; Search time 241 Seconds

(Without alignments)  
1593.148 Million cell updates/sec

Title: US-10-031-158B-13

Perfect score: 1027  
Sequence: 1 gggcagagtggtggcaaaaa.....atttaaaaaatgaaagct 1027

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3392430 seqs, 186927314 residues

Total number of hits satisfying chosen parameters: 6784860

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA.New:\*  
1: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/PCR\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq2:\*  
9: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq3:\*  
10: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	36.8	3.6	600	US-10-750-185-1996	Sequence 1996, Ap
C 2	34.6	3.4	1561	US-10-750-185-40906	Sequence 40906, A
C 3	34	3.3	905	US-10-750-185-53770	Sequence 53770, A
C 4	34	3.3	1173	US-10-750-185-52322	Sequence 52322, A
C 5	34	3.3	110837	US-11-121-086-87	Sequence 87, Appl
C 6	34	3.3	191797	US-11-121-086-13	Sequence 13, Appl
C 7	33.8	3.3	96988	US-11-117-187-196	Sequence 156, Appl
C 8	33.6	3.3	182303	US-11-121-086-45	Sequence 45, Appl
C 9	33.4	3.3	1515	US-10-750-185-40420	Sequence 40420, A
C 10	33.4	3.3	1567	US-10-750-185-25139	Sequence 25139, A
C 11	33.2	3.2	40439	US-10-993-509-1	Sequence 1, Appl
C 12	33	3.2	1128	US-10-750-185-54062	Sequence 54062, A
C 13	33	3.2	108214	US-11-117-187-211	Sequence 211, App
C 14	32.8	3.2	600	US-10-750-185-102	Sequence 102, Appl
C 15	32.8	3.2	1037	US-10-518-753-14	Sequence 14, Appl
C 16	32.8	3.2	1449	US-10-750-185-53517	Sequence 53517, A
C 17	32.8	3.2	2767	US-10-750-185-24844	Sequence 24844, A
C 18	32.8	3.2	4497	US-10-518-753-15	Sequence 15, Appl
C 19	32.8	3.2	162173	US-11-121-086-72	Sequence 72, Appl
C 20	32.4	3.2	161874	US-11-121-086-75	Sequence 75, Appl
C 21	32.2	3.1	600	US-10-750-185-20145	Sequence 20145, A
C 22	32.2	3.1	944	US-10-750-185-39148	Sequence 39148, A
C 23	32.2	3.1	2069	US-10-750-185-49897	Sequence 49897, A

24	32.2	3.1	186442	US-11-121-086-104	Sequence 104, App
C 25	32	3.1	1707	US-10-750-185-40279	Sequence 40279, A
C 26	31.8	3.1	4265	US-11-044-051-74	Sequence 74, Appl
C 27	31.8	3.1	260209	US-10-933-023-23	Sequence 23, Appl
C 28	31.6	3.1	319608	US-11-145-703-1	Sequence 1, Appl
C 29	31.4	3.1	319	US-11-108-172-339	Sequence 339, App
C 30	31.4	3.1	1096	US-10-750-185-25016	Sequence 25016, A
C 31	31.4	3.1	1441	US-10-750-185-48298	Sequence 48298, A
C 32	31.4	3.1	1564	US-10-750-185-48597	Sequence 48597, A
C 33	31.4	3.1	1838	US-10-750-185-51829	Sequence 51829, A
C 34	31.2	3.0	1034	US-10-750-185-40337	Sequence 40337, A
C 35	31.2	3.0	2744	US-10-750-185-30635	Sequence 30635, A
C 36	31.2	3.0	3155	US-10-750-185-59681	Sequence 59681, A
C 37	31.2	3.0	176760	US-11-121-086-51	Sequence 51, Appl
C 38	31	3.0	817	US-10-750-185-53425	Sequence 53425, A
C 39	31	3.0	1077	US-10-750-185-51726	Sequence 51726, A
C 40	31	3.0	3755	US-10-793-626-3642	Sequence 3642, Ap
C 41	31	3.0	4070	US-10-750-185-53698	Sequence 53698, A
C 42	30.8	3.0	154548	US-11-121-086-33	Sequence 33, Appl
C 43	30.8	3.0	1025	US-10-750-185-44629	Sequence 44629, A
C 44	30.8	3.0	2511	US-10-750-185-41003	Sequence 41003, A
C 45	30.8	3.0	3272	US-10-750-185-24646	Sequence 24646, A

#### ALIGNMENTS

RESULT 1  
US-10-750-185-1996/c  
; Sequence 1996, Application US/10750185  
; Publication No. US2005026603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMT GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1996  
; LENGTH: 600  
; TYPE: DNA  
; ORGANISM: Bovine MM118519  
; US-10-750-185-1996

Query Match 3.6%; Score 36.8; DB 6; Length 600;  
Best Local Similarity 52.0%; Pred. No. 0.17;  
Matches 106; Conservative 0; Mismatches 97; Indels 1; Gaps 1;

163 GAGAAATTTTCCCTGATTTAATGATCATTTGGCAAGAAAGACACAGATT 222  
Db 528 GATTAATTTTCAAGCTATATGATTAAGCCCTTCTGTTAAAGATATGTTATGTTT 469  
QY 223 CTGGATCCAGAGGGAACACATGACATTAAGCAACATTAAGCAATTAAGTCTGG 282  
QY 468 ATAGGTCTATTATTAAGAAAAATCATACACACATCAAAACCAAGATCTTATGACA 409  
Db 283 TTA-ACGGTCCAGAAAAGTCATGACAAAGACAGATGTATCGTCAGACATGAGAA 341  
Db 408 TTATCTCTTGAATTAAGATTAACGTTTAAAGAAAAAAGTGAATGATTAAGAGAAA 349  
QY 342 TTAATAAACGAGATTGATCAAGA 365  
Db 348 TATTAAGACGTAAACGTTCAAGA 325

RESULT 2  
US-10-750-185-40906/c  
; Sequence 40906, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; PRIOR FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 40906  
; LENGTH: 1561  
; TYPE: DNA  
; ORGANISM: Bovine 19866881662064  
US-10-750-185-40906

Query Match 3.4%; Score 34.6; DB 6; Length 1561;  
Best Local Similarity 61.8%; Pred. No. 1.5;

Matches 55; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 655 TTGGCTTTCTTTCTTGGGTTGGGCAATTCAGTTCTGATGTGTACTATTATCAATT 714  
Db 1426 TTGGCCATGTCACATGATGATGGAGCTTAGTTCCTGCTGTATATCACTGAGG 1367  
Qy 715 ATTGATTAACGTTTTCACCAACGAGGCG 743  
Db 1366 TGTATGACTCTTGTGCACTCTGAGGC 1338

RESULT 3  
US-10-750-185-53770  
; Sequence 53770, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; PRIOR FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 53770  
; LENGTH: 905  
; TYPE: DNA  
; ORGANISM: Bovine 19866880693972  
US-10-750-185-53770

Query Match 3.3%; Score 34; DB 6; Length 905;  
Best Local Similarity 50.0%; Pred. No. 1.6;  
Matches 85; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

Qy 855 ATAGGTAGACATCTCTGGGCTTACGCTTCCCTCTTACGTCTTATCAAT 914  
Db 456 ATAGTACATCAATCAAGGTGAGGCTGCTCTCTCACTGGATATATGATGA 515

Qy 915 ACTGCTGAGACCTTTCATTTTACAGCCCTGAGACATCTTCTTGTAGTTGAAT 974  
Db 516 CATTCGCTTCCCTCTACGTGTTGCTGTGTATTTCTATTTTCATATAAAC 575  
Qy 975 ATGTGTTGTTTTCCTGTAATAGCAAAATTAATTTAAAAATGAAA 1024  
Db 576 ATGTTAATTTTATTAATAGAAAAAATTTTAAATTAATAAAAAA 625

RESULT 4  
US-10-750-185-52322/c  
; Sequence 52322, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; PRIOR FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 52322  
; LENGTH: 1173  
; TYPE: DNA  
; ORGANISM: Bovine 19866880907657  
US-10-750-185-52322

Query Match 3.3%; Score 34; DB 6; Length 1173;  
Best Local Similarity 66.2%; Pred. No. 1.9;  
Matches 49; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 953 AGCTCTTGTCTGTAATTAATGATGTTTCCGTAATACAAATAATTA 1012  
Db 231 ATTAGTATTTACATGTAATTTGATGAGATGTTGCTTACACATACATTTA 172  
Qy 1013 AAAAATGAAAAAT 1026  
Db 171 CCAAAAAATTAAT 158

RESULT 5  
US-11-121-086-97/c  
; Sequence 97, Application US/11121086  
; Publication No. US20050266459A1  
; GENERAL INFORMATION:  
; APPLICANT: FOUJSEN, TIM S.  
; APPLICANT: NIELSEN, KIRSTEN V.  
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES  
; FILE REFERENCE: 09138.6000-00000  
; CURRENT APPLICATION NUMBER: US/11/121,086  
; PRIOR FILING DATE: 2005-05-04  
; PRIOR APPLICATION NUMBER: 60/567,570  
; PRIOR FILING DATE: 2004-05-04  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 97  
; LENGTH: 170837  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-121-086-97

Query Match 3.3%; Score 34; DB 7; Length 170837;  
Best Local Similarity 52.1%; Pred. No. 40;  
Matches 76; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

RESULT 9  
US-10-750-185-40420  
; Sequence 40420, Application US/10750185  
; Publication No. US20050260603A1

```

; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40420
; LENGTH: 1515
; TYPE: DNA
; ORGANISM: Bovine 19866881526906
US-10-750-185-40420
```

```

Query Match      3.3% Score 33.4; DB 6; Length 1515;
Best Local Similarity 46.7%; Pred. No. 3.4;
Matches 106; Conservative 0; Mismatches 121; Indels 0; Gaps 0;
```

```

Qy 25 AAGGATTTGGTCCCGAACAAGCTTATCATTAAGATAAACAATTGATGAGATGTT 84
Db 23 AATCAGATTTCAGCAAAATAGATATTAATTAGTCAGACAGGACAAACCTTTT 82
Qy 85 TCCCCCAAGCCACTATTTTCTCTTCAATTGCTGAACAAGCTCCAGAGCTGGA 144
Db 83 TCCCAACAACACTGGAACCTGCTTCTGTCTTTAGGAACCTGCAAGGACCTT 142
Qy 145 ACATACCTTCTCTCTGGAAGATTTTCCCTGATGTTATTAAGATACATTGGCAAGAA 204
Db 143 ACATGAATCATTCATTTTGTAAATACAAATTTCTGACACAGCGTAGCTAGCCAGT 202
Qy 205 AAGAAGACAACAGATTCTGGATCCGAGAGGGAACACCATGAA 251
Db 203 TCGATTGAGAAACAGATCAGCATTTAATGATGATGACACAGATGAA 249
```

```

RESULT 10
US-10-750-185-25199/c
; Sequence 25199, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25199
; LENGTH: 1567
; TYPE: DNA
; ORGANISM: Bovine 19866880415063
US-10-750-185-25199
```

```

Query Match      3.3% Score 33.4; DB 6; Length 1567;
Best Local Similarity 47.8%; Pred. No. 3.4;
Matches 97; Conservative 0; Mismatches 106; Indels 0; Gaps 0;
```

```

Qy 283 TTAAGTGGCCAGAAAAGTCATGACAAAGAACACAGATGTATCTGCAGACATGAGAT 342
Db 582 TTATGCTTGAAGAAATTAATTTATTTAGAGCAACAAGCATGTGACCTTGTAAT 523
Qy 343 AATATAAACGAGTGTGATCAAGAAATATATCTTCTCCATTAAGACGATGTGCACA 402
Db 522 TATCTCAATGAGAGTATCTCAGAAAGATAGTAGCACACATATGACTTATTCACATG 463
Qy 403 ATGATCCCAAGACAATTTGTCAAAAGATGCAATGATACACTACTGCTGCAGCTACA 462
Db 462 CAATGGCACCAGATCATCTCGAAGACATGACAGATATTCAGAGCAGATCATTTAAAAA 403
Qy 463 AACACCTCTGCATTTATCATGTA 485
Db 402 AATTTTCTATTTAGCTTGAA 380
```

```

RESULT 11
US-10-993-509-1
; Sequence 1, Application US/10993509
; Publication No. US20050250121A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Aetna, Jercen
; APPLICANT: Athansiou, Maria
; APPLICANT: Brain, Carlos
; APPLICANT: Cohen, Nadine
; APPLICANT: Dain, Bradley
; APPLICANT: Denton, R. Rex
; APPLICANT: Judson, Richard S.
; APPLICANT: Ozdemir, Vural
; APPLICANT: Reed, Carol R.
; TITLE OF INVENTION: NTRK Genetic Markers Associated with Progression of Alzheimer's
; FILE REFERENCE: 2300.0060001
; CURRENT APPLICATION NUMBER: US/10/993,509
; PRIOR FILING DATE: 2004-11-22
; PRIOR APPLICATION NUMBER: US 60/524,637
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 40439
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

```

NAME/KEY: misc feature
LOCATION: (2402)..(2402)
OTHER INFORMATION: n is 'c' or 'g'
FEATURE:
NAME/KEY: misc feature
LOCATION: (2722)..(2722)
OTHER INFORMATION: n is 'c' or 't'
FEATURE:
NAME/KEY: misc feature
LOCATION: (2799)..(2799)
OTHER INFORMATION: n is 'g' or 'a'
FEATURE:
NAME/KEY: misc feature
LOCATION: (3785)..(3898)
OTHER INFORMATION: n is 'a', 't', 'g', or 'c'
FEATURE:
NAME/KEY: misc feature
LOCATION: (6337)..(6436)
OTHER INFORMATION: n is 'a', 't', 'g', or 'c'
FEATURE:
NAME/KEY: misc feature
LOCATION: (11729)..(11828)
OTHER INFORMATION: n is 'a', 't', 'g', or 'c'
FEATURE:
NAME/KEY: misc feature
LOCATION: (14905)..(15004)
OTHER INFORMATION: n is 'a', 't', 'g', or 'c'
FEATURE:
```

```

; NAME/KEY: misc feature
; LOCATION: (117523)..(117622)
; OTHER INFORMATION: n is 'a', 'c', 'g', or 'c'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (119442)..(119441)
; OTHER INFORMATION: n is 'a', 'c', 'g', or 'c'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (122316)..(122415)
; OTHER INFORMATION: n is 'a', 'c', 'g', or 'c'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (124686)..(124785)
; OTHER INFORMATION: n is 'a', 'c', 'g', or 'c'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (127030)..(127129)
; OTHER INFORMATION: n is 'a', 'c', 'g', or 'c'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (129535)..(129634)
; OTHER INFORMATION: n is 'a', 'c', 'g', or 'c'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (131929)..(132028)
; OTHER INFORMATION: n is 'a', 'c', 'g', or 'c'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (134403)..(134502)
; OTHER INFORMATION: n is 'a', 'c', 'g', or 'c'
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (136929)..(137028)
; OTHER INFORMATION: n is 'a', 'c', 'g', or 'c'
; US-10-993-509-1

```

Query Match 3.2%; Score 33.2; DB 6; Length 40439;

Best Local Similarity 34.5%; Pred. No. 29; Matches 116; Conservative 0; Mismatches 218; Indels 2; Gaps 1;

```

Qy 485 ACCTCTCTCTCTCTCTCAAGAGTGTCTATTTTGGCATCATCACTGCTGTCTGTTA 544
Db 34298 ACTGCAACCTCGCGCTCCAGTTCAAGTATCTCCACCTCACTCCAGAGTGTG 34357
Qy 545 GAAGAACGGCTTCTCTGTCGATGAGAGAAATCATPAACAGCGTGGCAAGAGGCC 604
Db 34358 GATTAACAGGATGTGCGACCATGTCTGTAT--TTTCATATTTTNNNNNNNNNN 34415
Qy 605 ATCTTTTCTCATCGGTATGTGCTCAGAGCGTCTTGAAGATCTAGTGGCTTTC 664
Db 34416 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 34475
Qy 665 TTCTGGGTTTGGGCACTTCAGTCTCATGTGTACTATTCATCATTTGTTAAC 724
Db 34476 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 34535
Qy 725 GGTTCCTCAACCAAGTGGGCAACAGAACCTCATCTGTATPAATCAATGAGAAATAGCC 764
Db 34536 TGTTCCTCAACCAAGTGGGCAACAGAACCTCATCTGTATPAATCAATGAGAAATAGCC 34595
Qy 785 ACGGCGATCTCCAGCAACCAATCTCTCATGTTTTC 820
Db 34596 ACGGCGAGCCCGACAGCAACCCCGCTTATTTCC 34631

```

RESULT 12

```

US-10-750-185-54062
; Sequence 54062, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.

```

```

; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54062
; LENGTH: 1128
; TYPE: DNA
; ORGANISM: Bovine
; US-10-750-185-54062

```

Query Match 3.2%; Score 33; DB 6; Length 1128;

Best Local Similarity 52.6%; Pred. No. 3.7; Matches 72; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

```

Qy 274 TTTAGCTGTAAACGTTGCGAGAAAGTCACTGACAAAGAACAGATGTATGTCAGA 333
Db 708 TTTAAGTGAAGTGAAGTGGGAATTAATCACTAGTATTTACACATTTTAAACATTTGA 767
Qy 334 CATGACATTAATTAACGAGGATGATCAAGAAATATCTTCTCCATTAAGAGGAT 393
Db 768 CATATTTAAAGATCAAGTACCTGTTTAAATGTTGCTCAAAATTTACAAAG 827
Qy 394 GTCATCAATGATGATCC 410
Db 828 CAGAGCAATTTATATTC 844

```

RESULT 13

```

US-11-117-187-211
; Sequence 211, Application US/11117187
; Publication No. US2005026560A1
; GENERAL INFORMATION:
; APPLICANT: PREUS, DAPHNE
; APPLICANT: COPEHAYER, GREGORY
; TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS
; FILE REFERENCE: ARCD:309US
; CURRENT APPLICATION NUMBER: US/11/117,187
; CURRENT FILING DATE: 2005-04-28
; PRIOR APPLICATION NUMBER: US/09/531,120
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/125,219
; PRIOR FILING DATE: 1999-03-18
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 211
; LENGTH: 1082144
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-11-117-187-211

```

Query Match 3.2%; Score 33; DB 7; Length 1082144;

Best Local Similarity 43.5%; Pred. No. 1.8e+02; Matches 150; Conservative 0; Mismatches 195; Indels 0; Gaps 0;

```

Qy 159 TCTTGAAATTTTCCCTGATGTTATTAAGATACATTGGCAAGAAAGAGCAACAC 218
Db 493328 TATTTAAATATTAATGTTATTTATTTATTAATTTTAAACCAAAATATTAATTAACAC 493387
Qy 219 GATTCGGAATCCAGAGAGGGAACACATGAAGATACACATATGATGAATTTG 278
Db 493388 AATTAATCTAATAATCTTTGATGATTCATTTGTTTATGTTATTTATTAATTAATAC 493447
Qy 279 CTGTTAAGGTCGCAAGAAAGTCACTGCAAGAAACAGATGTATGTCAGACATGA 338

```

Db 493448 AAGTTGCAAACTTAAACCAATTTTAAATAAGTAAACGATTTATTTAAATTAATAA 493507

QY 339 GAATATATAAAAACGAGTTTGATCAGAGAAATTATCTTTCTCCATATAAGACGATGTCAAT 398

Db 493508 CAATTTTATAGTCTTTTCTTTTCTTCCAAAAACGTGTGCGAAAACCGATCTGAATCCATATTCAT 493567

QY 399 CACATGTGATCCCAAGAACAAATTTGTTCAAAAGATGAATGTATACACTAGCTGTGAGCT 458

Db 493568 CCAAAACCGAAACCAAACTGTATATGTATATTAATCCGACCGGATTAATTAACCTTTCATCC 493627

QY 459 CACAACAACACTCTGCATATTAACATGACCTCCCTCGTCTCTCA 503

Db 493628 AATTTATGCAATTCAAAAAAATTGCATGTGATTCGAAATCCGAAA 493672

```

RESULT 14
US-10-750-185-102/c
, Sequence 102, Application US/10750185
, Publication No. US20050260603A1
, GENERAL INFORMATION:
, APPLICANT: MMI GENOMICS, INC.
, APPLICANT: DENISE, Sue K.
, APPLICANT: KERR, Richard
, APPLICANT: ROSENFELD, David
, APPLICANT: HOLM, Tom
, APPLICANT: BATES, Stephen
, APPLICANT: FANTIN, Dennis
, TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
, FILE REFERENCE: MM1100-2
, CURRENT APPLICATION NUMBER: US/10/750,185
, CURRENT FILING DATE: 2003-12-31
, PRIOR APPLICATION NUMBER: US 60/437,482
, PRIOR FILING DATE: 2002-12-31
, NUMBER OF SEQ ID NOS: 64922
, SOFTWARE: PatentIn version 3.1
, SEQ ID NO 102
, LENGTH: 600
, TYPE: DNA
, ORGANISM: Bovine
MMBT02133
US-10-750-185-102

```

	Query Match	3.2%	Score 32.8	DB 6	Length 600
	Best Local Similarity	59.8%	Pred. No. 2.9		
	Matches	55	Conservative	0	Mismatches 37; Indels 0; Gaps 0;
Oy	18 AAAAAATCAGGATTGTGGTCCCGGAACAAGGTTATCATTCACATAAACAACCTGATGC	77			
Dd	193 AAAAAATGAAAATATAGATCCCTGTCCAAGATATATTAAGAATTTCAAGAAAGTGACAGC	134			
Oy	78 AGATGTTTCCCCCAAGCCCACTATTTTTCTTC	109			
Dd	133 AGAGTATTACACCAGCAACAGGCCCTTTTCTGC	102			

```

RESULT 15
US-10-518-753-14
; Sequence 14, Application US/10518753
; Publication No. US20050262589A1
; GENERAL INFORMATION:
; APPLICANT: Calgene LLC
; TITLE OF INVENTION: Nucleic Acid Sequences and Methods of Use for the Production of Fatty Acids
; TITLE OF INVENTION: with Modified Polyunsaturated Fatty Acids
; FILE REFERENCE: 16518.129
; CURRENT APPLICATION NUMBER: US/10/518,753
; CURRENT FILING DATE: 2004-12-21
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 1037
; TYPE: DNA
; ORGANISM: Glycine max
US-10-518-753-14

```

```

Query Match          3.2%; Score 32.8; DB 6; Length 1037;
Best Local Similarity 59.8%; Pred. No. 4.1;
Matches 55; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY      136 AAGCGTGGAAACATACCTTTGCTCTTGAGAAATTTTCCCTGATGTTATTAGATACAT 195
DB      39 AATGTAGATACACCTCTCTCTCTTTTCTTTTCTTTTGTGAGGTTAATGCTAGATAA 98
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      196 TGGCAAGAAAAAGAAAGACACAGATTCTGGG 227
DB      99 TAGCTAGAAAGAGAAAGAACACAAATTAGG 130
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Search completed: December 11, 2005, 01:35:02
Job time : 245 secs

```

```
Search completed: December 11, 2005, 01:35:02
Job time : 245 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2005, 22:15:03 ; Search time 890 Seconds

(without alignments)  
9542.305 Million cell updates/sec

Title: US-10-031-158b-13

Perfect score: 1027

Sequence: 1 GGGCAAGAGTGGGCAAAA.....atttaaaataatgaagaatc 1027

Scoring table: IDENTITY\_NUC

Gap 10.0 , Gape 1.0

Searched: 9793542 seqs, 413468905 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA Main:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1027	100.0	1027	5	US-10-205-823-404
2	1027	100.0	1027	5	US-10-756-149-2859
3	1027	100.0	1027	10	US-11-051-454-404
4	1013.8	98.7	1395	8	US-10-643-795A-23
5	1013.8	98.7	1395	6	US-10-948-518-23
6	968.2	94.3	1799	6	US-10-101-510-492
7	945.6	92.1	1418	3	US-09-957-708-16
8	906	88.2	1155	3	US-09-925-300-58
9	902.2	87.8	1421	3	US-09-954-456-317
10	902.2	87.8	1421	3	US-10-843-641A-3344
11	838	81.6	1586	3	US-09-960-706-676
12	838	81.6	1586	3	US-09-873-319-424
13	838	81.6	1586	9	US-10-847-918-14
14	795.8	77.5	1162	8	US-10-357-930-21517
15	795.8	77.5	1162	8	US-10-357-930-24142
16	795.8	77.5	1162	8	US-10-357-930-24529
17	795.8	77.5	1162	8	US-10-357-930-24533
18	795.8	77.5	1162	8	US-10-357-930-24840
19	795.8	77.5	1162	8	US-10-357-930-25151
20	795.8	77.5	1162	8	US-10-357-930-27359
21	795.8	77.5	1162	8	US-10-357-930-27740
22	619	60.3	924	8	US-10-357-930-25081
23	589	57.4	124933	9	US-10-737-082-97

24	589	57.4	124933	9	US-10-765-790-97	Sequence 97, Appl
C 25	498.8	48.6	1316	8	US-10-357-930-25082	Sequence 25082, A
C 26	498.8	48.6	1316	8	US-10-357-930-26474	Sequence 26474, A
C 27	473.2	46.1	486	8	US-10-357-930-25002	Sequence 25002, A
C 28	449.2	43.7	825	6	US-10-340-536-29	Sequence 29, Appl
C 29	446.8	43.5	784	8	US-10-357-930-21887	Sequence 21887, A
C 30	444	43.2	511	3	US-09-759-143-72	Sequence 72, Appl
C 31	444	43.2	511	3	US-09-780-669-72	Sequence 72, Appl
C 32	444	43.2	511	3	US-09-030-606-72	Sequence 72, Appl
C 33	444	43.2	511	3	US-09-822-827-72	Sequence 72, Appl
C 34	444	43.2	511	3	US-09-115-453-72	Sequence 72, Appl
C 35	444	43.2	511	3	US-09-232-880-72	Sequence 72, Appl
C 36	444	43.2	511	3	US-09-895-793-72	Sequence 72, Appl
C 37	444	43.2	511	3	US-09-895-814-72	Sequence 72, Appl
C 38	444	43.2	511	5	US-10-012-896-72	Sequence 72, Appl
C 39	444	43.2	511	5	US-10-010-940-72	Sequence 72, Appl
C 40	444	43.2	511	6	US-10-144-678A-72	Sequence 72, Appl
C 41	444	43.2	511	6	US-10-294-025-72	Sequence 72, Appl
C 42	444	43.2	511	7	US-10-688-838-72	Sequence 72, Appl
C 43	416.2	40.5	569	8	US-10-357-930-43494	Sequence 43494, A
C 44	410.4	40.0	539	8	US-10-357-930-22784	Sequence 22784, A
C 45	410.4	40.0	539	8	US-10-357-930-28636	Sequence 28636, A

#### ALIGNMENTS

RESULT 1  
US-10-205-823-404  
Sequence 404, Application US/10205823  
Publication No. US20030108963A1  
GENERAL INFORMATION:  
APPLICANT: Schlegel, Robert  
APPLICANT: Monahan, John E.  
APPLICANT: Endege, Wilson O.  
APPLICANT: Ganavarepu, Manjula  
APPLICANT: Gorbatcheva, Bella  
APPLICANT: Hoersch, Sebastian  
APPLICANT: Kametkar, Shubhangi  
APPLICANT: Monsey, Angela M.  
APPLICANT: Glatc, Karen  
APPLICANT: Anderson, Dustin  
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF PROSTATE CANCER  
FILE REFERENCE: MRI-04  
CURRENT FILING DATE: 2002-07-25  
PRIOR APPLICATION NUMBER: US/10/205, 823  
PRIOR FILING DATE: 2001-07-25  
PRIOR APPLICATION NUMBER: 60/314,356  
PRIOR FILING DATE: 2001-08-22  
PRIOR APPLICATION NUMBER: 60/325,020  
PRIOR FILING DATE: 2001-09-25  
PRIOR APPLICATION NUMBER: 60/341,746  
PRIOR FILING DATE: 2001-12-12  
PRIOR APPLICATION NUMBER: 60/362,158  
PRIOR FILING DATE: 2002-03-05  
NUMBER OF SEQ ID NOS: 455  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 404  
LENGTH: 1027  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-205-823-404  
Query Match 100.0%; Score 1027; DB 5; Length 1027;  
Best Local Similarity 100.0%; Pred. No. 1.1e-293;  
Matches 1027; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 GGGCAAGTGGGCAAAAATCAAGTATTTGTCGCGAACAAAGCTTATCATTTACA 60  
|||||

Db 1 GGGCAAGATTGGGCAAAAAATCAAGTATTGGTCCGGACAAAGCTTATCATTAACA 60  
Qy 61 GATPAAACAATTGATGAGATGTTTCCCAAGCCCATATTTTCTTCTTCAATTGCT 120  
Db 61 GATPAAACAATTGATGAGATGTTTCCCAAGCCCATATTTTCTTCTTCAATTGCT 120  
Qy 121 GAAACAAAGCTCCAGAAAGCTGGAAACATACCTTTGTCTTCTTGAAGAAATTTTCCCTGAT 180  
Db 121 GAAACAAAGCTCCAGAAAGCTGGAAACATACCTTTGTCTTCTTGAAGAAATTTTCCCTGAT 180  
Qy 181 GTTATTAAAGATTCATGGCAAGAAAGAAAGCAACAGATTCCTGGATCCCAAGAGGG 240  
Db 181 GTTATTAAAGATTCATGGCAAGAAAGAAAGCAACAGATTCCTGGATCCCAAGAGGG 240  
Qy 241 AACACATGAAGACTTAACGACATACATCAATTAATTTAGCTGGTTAACGGTCCAGAAAAAG 300  
Db 241 AACACATGAAGACTTAACGACATACATCAATTAATTTAGCTGGTTAACGGTCCAGAAAAAG 300  
Qy 301 TCACCTGACAAAGAACACAGATGTATCGTCAGACATGAGAAATTAATAAAAACGAGATTGAT 360  
Db 301 TCACCTGACAAAGAACACAGATGTATCGTCAGACATGAGAAATTAATAAAAACGAGATTGAT 360  
Qy 361 CAAGAAATTAATCTTCTCCCAATTAAGACGATGTCAACAAACACCTCTGCATATTAC 420  
Db 361 CAAGAAATTAATCTTCTCCCAATTAAGACGATGTCAACAAACACCTCTGCATATTAC 420  
Qy 421 TGTTCAAAAGATGCAAAATGATACACTACTGCTGACAGTCAAGAAACACCTCTGCATATTAC 480  
Db 421 TGTTCAAAAGATGCAAAATGATACACTACTGCTGACAGTCAAGAAACACCTCTGCATATTAC 480  
Qy 481 ATGTACCTCTCTGCTGCTCCAGAGTGTGTCTATTTTGGCAATCATCACTGCTGTCTG 540  
Db 481 ATGTACCTCTCTGCTGCTCCAGAGTGTGTCTATTTTGGCAATCATCACTGCTGTCTG 540  
Qy 541 CTTCGAAAAGAGCGCTTTTCTGCTGCAATGAGAGAAATCATTAACAGCGTGGCAAGAGA 600  
Db 541 CTTCGAAAAGAGCGCTTTTCTGCTGCAATGAGAGAAATCATTAACAGCGTGGCAAGAGA 600  
Qy 601 GGCACATCTTTCTCATCGGTTATATGTCCTTAGAAGCGTCTTCTGAGGATCTAGTTGGGC 660  
Db 601 GGCACATCTTTCTCATCGGTTATATGTCCTTAGAAGCGTCTTCTGAGGATCTAGTTGGGC 660  
Qy 661 TTTCTTTCTGGGTTTGGGCATTTTCAGTTCTCATGTGTGTAATACTATTATTAATTGTA 720  
Db 661 TTTCTTTCTGGGTTTGGGCATTTTCAGTTCTCATGTGTGTAATACTATTATTAATTGTA 720  
Qy 721 TAAAGGTTTTCAAAACCAATGAGGACACAGAGAACTCTACTCTGTAAATAACAATAGAGAT 780  
Db 721 TAAAGGTTTTCAAAACCAATGAGGACACAGAGAACTCTACTGTAAATAACAATAGAGAT 780  
Qy 781 AGCCACGGCGATCTCCAGACCAATCTCTCCATGTTTTCACAGCTCTCTCCAGCCCAACC 840  
Db 781 AGCCACGGCGATCTCCAGACCAATCTCTCCATGTTTTCACAGCTCTCTCCAGCCCAACC 840  
Qy 841 AAATAGCGCTGCTATGATGTAAGATCTCGGCGCTTTCAGCTTTCCTCTTAAATG 900  
Db 841 AAATAGCGCTGCTATGATGTAAGATCTCGGCGCTTTCAGCTTTCCTCTTAAATG 900  
Qy 901 TTTCTTTATTCAGATACTGCTCGGAAAGCTTTTCAATTTTACAAGCCCTGAAGAGCTTCT 960  
Db 901 TTTCTTTATTCAGATACTGCTCGGAAAGCTTTTCAATTTTACAAGCCCTGAAGAGCTTCT 960  
Qy 961 TTGCTAGTTGAATTAATGATGATGTTTTCCTGTAATAGCAAAATTAATTTAAAAAATG 1020  
Db 961 TTGCTAGTTGAATTAATGATGATGTTTTCCTGTAATAGCAAAATTAATTTAAAAAATG 1020  
Qy 1021 AAAAGTT 1027  
Db 1021 AAAAGTT 1027

/ Sequence 2859, Application US/10756149  
/ Publication No. US20050181375A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Aziz, Nafae  
/ APPLICANT: Zlotnik, Albert  
/ TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND  
/ FILE REFERENCE: file  
/ CURRENT APPLICATION NUMBER: US/10/756,149  
/ CURRENT FILING DATE: 2004-01-12  
/ NUMBER OF SEQ ID NOS: 5818  
/ SOFTWARE: Patentin version 3.2  
/ SEQ ID NO 2859  
/ LENGTH: 1027  
/ TYPE: DNA  
/ ORGANISM: Homo Sapiens  
US-10-756-149-2859

Query Match 100.0%; Score 1027; DB 9; Length 1027;  
Best Local Similarity 100.0%; Pred. No. 1.1e-293;  
Matches 1027; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGCAAGATTGGGCAAAAAATCAAGTATTGGTCCGGACAAAGCTTATCATTAACA 60  
Db 1 GGGCAAGATTGGGCAAAAAATCAAGTATTGGTCCGGACAAAGCTTATCATTAACA 60  
Qy 61 GATPAAACAATTGATGAGATGTTTCCCAAGCCCATATTTTCTTCTTCAATTGCT 120  
Db 61 GATPAAACAATTGATGAGATGTTTCCCAAGCCCATATTTTCTTCTTCAATTGCT 120  
Qy 121 GAAACAAAGCTCCAGAAAGCTGGAAACATACCTTTGTCTTCTTGAAGAAATTTTCCCTGAT 180  
Db 121 GAAACAAAGCTCCAGAAAGCTGGAAACATACCTTTGTCTTCTTGAAGAAATTTTCCCTGAT 180  
Qy 181 GTTATTAAAGATTCATGGCAAGAAAGAAAGCAACAGATTCCTGGATCCCAAGAGGG 240  
Db 181 GTTATTAAAGATTCATGGCAAGAAAGAAAGCAACAGATTCCTGGATCCCAAGAGGG 240  
Qy 241 AACACATGAAGACTTAACGACATACATCAATTAATTTAGCTGGTTAACGGTCCAGAAAAAG 300  
Db 241 AACACATGAAGACTTAACGACATACATCAATTAATTTAGCTGGTTAACGGTCCAGAAAAAG 300  
Qy 301 TCACCTGACAAAGAACACAGATGTATCGTCAGACATGAGAAATTAATAAAAACGAGATTGAT 360  
Db 301 TCACCTGACAAAGAACACAGATGTATCGTCAGACATGAGAAATTAATAAAAACGAGATTGAT 360  
Qy 361 CAAGAAATTAATCTTCTCCCAATTAAGACGATGTCAACAAACACCTCTGCATATTAC 420  
Db 361 CAAGAAATTAATCTTCTCCCAATTAAGACGATGTCAACAAACACCTCTGCATATTAC 420  
Qy 421 TGTTCAAAAGATGCAAAATGATACACTACTGCTGACAGTCAAGAAACACCTCTGCATATTAC 480  
Db 421 TGTTCAAAAGATGCAAAATGATACACTACTGCTGACAGTCAAGAAACACCTCTGCATATTAC 480  
Qy 481 ATGTACCTCTCTGCTGCTCCAGAGTGTGTCTATTTTGGCAATCATCACTGCTGTCTG 540  
Db 481 ATGTACCTCTCTGCTGCTCCAGAGTGTGTCTATTTTGGCAATCATCACTGCTGTCTG 540  
Qy 541 CTTCGAAAAGAGCGCTTTTCTGCTGCAATGAGAGAAATCATTAACAGCGTGGCAAGAGA 600  
Db 541 CTTCGAAAAGAGCGCTTTTCTGCTGCAATGAGAGAAATCATTAACAGCGTGGCAAGAGA 600  
Qy 601 GGCACATCTTTCTCATCGGTTATATGTCCTTAGAAGCGTCTTCTGAGGATCTAGTTGGGC 660  
Db 601 GGCACATCTTTCTCATCGGTTATATGTCCTTAGAAGCGTCTTCTGAGGATCTAGTTGGGC 660  
Qy 661 TTTCTTTCTGGGTTTGGGCATTTTCAGTTCTCATGTGTGTAATACTATTATTAATTGTA 720  
Db 661 TTTCTTTCTGGGTTTGGGCATTTTCAGTTCTCATGTGTGTAATACTATTATTAATTGTA 720  
Qy 721 TAAAGGTTTTCAAAACCAATGAGGACACAGAGAACTCTACTCTGTAAATAACAATAGAGAT 780  
Db 721 TAAAGGTTTTCAAAACCAATGAGGACACAGAGAACTCTACTGTAAATAACAATAGAGAT 780

Qy 781 AGCAGCGGATCTCCAGCACAATCTCTCATGTTTTTCCAGCTCTCCAGCAACC 840  
Db 781 AGCAGCGGATCTCCAGCACAATCTCTCATGTTTTTCCAGCTCTCCAGCAACC 840  
Qy 841 AATAGCGCTGTATAGTATAGATCTGCGGCTTCTAGCTTGTCTCTTCTAGTG 900  
Db 841 AATAGCGCTGTATAGTATAGATCTGCGGCTTCTAGCTTGTCTCTTCTAGTG 900  
Qy 901 TTCTTTATAGATTAAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGT 960  
Db 901 TTCTTTATAGATTAAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGT 960  
Qy 961 TTCTTTATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGT 1020  
Db 961 TTCTTTATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGT 1020  
Qy 1021 AAAAGTT 1027  
Db 1021 AAAAGTT 1027

## RESULT 3

US-11-051-454-404  
; Sequence 404, Application US/11051454  
; Publication No. US20050191673A1

## GENERAL INFORMATION:

APPLICANT: Schlegel, Robert  
APPLICANT: Monahan, John E.  
APPLICANT: Endege, Wilson O.  
APPLICANT: Gannavarapu, Manjula  
APPLICANT: Gorbacheva, Bella  
APPLICANT: Hoersch, Sebastian  
APPLICANT: Kamatkar, Shubhangi  
APPLICANT: Monney, Angela M.  
APPLICANT: Giacit, Karen  
APPLICANT: Zhao, Xumei  
APPLICANT: Anderson, Dustin  
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND  
TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
TITLE OF INVENTION: THERAPY OF PROSTATE CANCER  
FILE REFERENCE: MRI-044  
CURRENT APPLICATION NUMBER: US/11/051,454  
CURRENT FILING DATE: 2005-02-04  
PRIOR APPLICATION NUMBER: US/10/205,823  
PRIOR FILING DATE: 2002-07-25  
PRIOR APPLICATION NUMBER: 60/307,982  
PRIOR FILING DATE: 2001-07-25  
PRIOR APPLICATION NUMBER: 60/314,356  
PRIOR FILING DATE: 2001-08-22  
PRIOR APPLICATION NUMBER: 60/325,020  
PRIOR FILING DATE: 2001-09-25  
PRIOR APPLICATION NUMBER: 60/341,746  
PRIOR FILING DATE: 2001-12-12  
PRIOR APPLICATION NUMBER: 60/362,158  
PRIOR FILING DATE: 2002-03-05  
NUMBER OF SEQ ID NOS: 455  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 404  
LENGTH: 1027  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-11-051-454-404

Query Match 100.0%; Score 1027; DB 10; Length 1027;  
Best Local Similarity 100.0%; Pred. No. 1,1e-293;  
Matches 1027; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGCAAGAGTTGGGCAAAAAAATCAAGTATTTGGTCCCGAACAAGCTTATATATACA 60  
Db 1 GGGCAAGAGTTGGGCAAAAAAATCAAGTATTTGGTCCCGAACAAGCTTATATATACA 60  
Qy 61 GATAAACAACCTTGATGAGATGTTTCCCGCAAGCCACTATTTTCTTCTCAATTGCT 120

Db 61 GATAAACAACCTTGATGAGATGTTTCCCGCAAGCCACTATTTTCTTCTCAATTGCT 120  
Qy 121 GAAACAACCTTGATGAGATGTTTCCCGCAAGCCACTATTTTCTTCTCAATTGCT 180  
Db 121 GAAACAACCTTGATGAGATGTTTCCCGCAAGCCACTATTTTCTTCTCAATTGCT 180  
Qy 181 GTTATTAAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGT 240  
Db 181 GTTATTAAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGT 240  
Qy 241 AACACCATGAGAGCTAAGACACATACATGAAATTTAGCTGTTAAAGGTCGCAAAAG 300  
Db 241 AACACCATGAGAGCTAAGACACATACATGAAATTTAGCTGTTAAAGGTCGCAAAAG 300  
Qy 301 TCATCTGAGCAAGAAACAACAGATGATATGCTGACATGAGATTAATTAACGAGTTGAT 360  
Db 301 TCATCTGAGCAAGAAACAACAGATGATATGCTGACATGAGATTAATTAACGAGTTGAT 360  
Qy 361 CAAGAAATTAATCTTCTGCAATTAAGAGAGATGTCATGCAATGATCCCAAGCAAT 420  
Db 361 CAAGAAATTAATCTTCTGCAATTAAGAGAGATGTCATGCAATGATCCCAAGCAAT 420  
Qy 421 TGTTCAAAAGATGCAATGATATACATGCTGAGCTGACAAACAAGCTTGCATATAC 480  
Db 421 TGTTCAAAAGATGCAATGATATACATGCTGAGCTGACAAACAAGCTTGCATATAC 480  
Qy 481 ATGTACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
Db 481 ATGTACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
Qy 541 CTTAGAAAGAACGGCTTCTGCTGCAATGAGAGAAATCAATACAGACGGTGGCAAGA 600  
Db 541 CTTAGAAAGAACGGCTTCTGCTGCAATGAGAGAAATCAATACAGACGGTGGCAAGA 600  
Qy 601 GGCATCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660  
Db 601 GGCATCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660  
Qy 661 TTTCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720  
Db 661 TTTCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720  
Qy 721 TAAAGTTTTCACACAGTGGGACACAGAGAACTCACTGTAATTAACATGAGGAAT 780  
Db 721 TAAAGTTTTCACACAGTGGGACACAGAGAACTCACTGTAATTAACATGAGGAAT 780  
Qy 781 AGCAGCGGATCTCCAGCACAATCTCTCATGTTTTTCCAGCTCTCCAGCAACC 840  
Db 781 AGCAGCGGATCTCCAGCACAATCTCTCATGTTTTTCCAGCTCTCCAGCAACC 840  
Qy 841 AATAGCGCTGTATAGTATAGATCTGCGGCTTCTAGCTTGTCTCTTCTAGTG 900  
Db 841 AATAGCGCTGTATAGTATAGATCTGCGGCTTCTAGCTTGTCTCTTCTAGTG 900  
Qy 901 TTCTTTATAGATTAAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGT 960  
Db 901 TTCTTTATAGATTAAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGT 960  
Qy 961 TTCTTTATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGT 1020  
Db 961 TTCTTTATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGTATAGT 1020  
Qy 1021 AAAAGTT 1027  
Db 1021 AAAAGTT 1027

## RESULT 4

US-10-643-795A-23  
; Sequence 23, Application US/10643795A  
; Publication No. US20040241703A1  
; GENERAL INFORMATION:

APPLICANT: FREDERIC J. DESAUVAGE  
APPLICANT: GRETCHEN FRANTZ  
APPLICANT: KENNETH J. HILLAN  
APPLICANT: PAUL POLAKIS  
APPLICANT: ANDREW POLSON  
APPLICANT: VICTORIA SMITH  
APPLICANT: SUSAN D. SPENCER  
APPLICANT: THOMAS D. WU  
APPLICANT: ZEMIN ZHANG  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND  
TREATMENT OF TUMOR  
FILE REFERENCE: P5026R1-US  
CURRENT APPLICATION NUMBER: US/10/643,795A  
PRIORITY FILING DATE: 2003-08-19  
PRIORITY APPLICATION NUMBER: US 60/404,809  
PRIORITY FILING DATE: 2002-08-19  
PRIORITY APPLICATION NUMBER: US 60/405,645  
PRIORITY FILING DATE: 2002-08-21  
PRIORITY APPLICATION NUMBER: US 60/413,192  
PRIORITY FILING DATE: 2002-09-23  
PRIORITY APPLICATION NUMBER: US 60/419,008  
PRIORITY FILING DATE: 2002-10-15  
PRIORITY APPLICATION NUMBER: US 60/426,847  
PRIORITY FILING DATE: 2002-11-15  
PRIORITY APPLICATION NUMBER: US 60/484,959  
PRIORITY FILING DATE: 2003-07-02  
NUMBER OF SEQ ID NOS: 158  
SEQ ID NO 23  
LENGTH: 1395  
TYPE: DNA  
ORGANISM: Homo sapien  
US-10-643-795A-23

Query Match 98.7%; Score 1013.8; DB 8; Length 1395;  
Best Local Similarity 99.8%; Pred. No. 1,1e-289;  
Matches 1015; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
4 CAAGAGTTGGGCAAAAAATCAAGTATTTGGTCCCGGACAAAGCTTATCATTAACAGAT 63  
379 CAAGAGTTGGGCAAAAAATCAAGTATTTGGTCCCGGACAAAGCTTATCATTAACAGAT 438  
64 AAACAACCTTGATGCAAGATGTTTCCCGCAAGCCCACTATTTTCTTCTTCAATGCTGAA 123  
439 AAACAACCTTGATGCAAGATGTTTCCCGCAAGCCCACTATTTTCTTCTTCAATGCTGAA 498  
124 ACAAGCTCCAGAGAGCTGGAACATACCTTTGCTTCTTGAAGAAATTTTCCCTGATGTT 183  
499 ACAAGCTCCAGAGAGCTGGAACATACCTTTGCTTCTTGAAGAAATTTTCCCTGATGTT 558  
184 ATTAAGATACATTTGGCAAGAAAGAGCAACAGATTTGGGATCCAGAGAGGGAAC 243  
559 ATTAAGATACATTTGGCAAGAAAGAGCAACAGATTTGGGATCCAGAGAGGGAAC 618  
244 ACCATGAAGACTTAAGCAACATCATGAAATTTAGCTGTTAAGCTGCCAGAAAAGTCA 303  
619 ACCATGAAGACTTAAGCAACATCATGAAATTTAGCTGTTAAGCTGCCAGAAAAGTCA 678  
304 CTGGAACAAGAACACAGATGATCGTCAGACATGAGATTAATAAAGAGGATGATCAA 363  
679 CTGGAACAAGAACACAGATGATCGTCAGACATGAGATTAATAAAGAGGATGATCAA 738  
364 GAAATTTATCTTCTCCATTAAGACGAGATGTCATCAATGATCCCAAGACATTTGT 423  
739 GAAATTTATCTTCTCCATTAAGACGAGATGTCATCAATGATCCCAAGACATTTGT 798  
424 TCAAAAGATGCAATGATGATCACTACTGCTGCAAGCTCAAAAACCTCTGCAATTAACATG 483  
799 TCAAAAGATGCAATGATGATCACTACTGCTGCAAGCTCAAAAACCTCTGCAATTAACATG 858  
484 TACCTCCCTGCTGCTCTCAAGAGTGTGATATTTTGGCATCATCACTGCTGCTGCTT 543  
859 TACCTCCCTGCTGCTCTCAAGAGTGTGATATTTTGGCATCATCACTGCTGCTGCTT 918

Qy 544 AGAAGAACGGCTTTCTGCTGCAATGAGAGAAATCATTAACAGACGGTGGCAAGAGGCC 603  
Db 919 AGAAGAACGGCTTTCTGCTGCAATGAGAGAAATCATTAACAGACGGTGGCAAGAGGCC 978  
Qy 604 CATCTTTCTCATCGGTTATTTGCTAGAGAGGCTTCTGAGAGATCTATGTTGGGCTT 663  
Db 979 CATCTTTCTCATCGGTTATTTGCTAGAGAGGCTTCTGAGAGATCTATGTTGGGCTT 1038  
Qy 664 CTTTCTGGGTTTGGGCAATTTCAATTTCTCATGTTGTTACTATTTCTATTTATTTATTA 723  
Db 1039 CTTTCTGGGTTTGGGCAATTTCAATTTCTCATGTTGTTACTATTTCTATTTATTTATTA 1098  
Qy 724 CGGTTTCAAAACCGATGGGCAACAGAGAACTTCACCTGTTAATTAACATGAGAAATAGC 783  
Db 1099 CGGTTTCAAAACCGATGGGCAACAGAGAACTTCACCTGTTAATTAACATGAGAAATAGC 1158  
Qy 784 CAGGGGATCTCCAGACCAATCTCTCATGTTTTCACAGAGCTCTCCAGCAACCCAAA 843  
Db 1159 CAGGGGATCTCCAGACCAATCTCTCATGTTTTCACAGAGCTCTCCAGCAACCCAAA 1218  
Qy 844 TAGGCTGCTATAGTATGATGATCATCTGCGGCTTTAGCTTGTCTCTTATGTTTC 903  
Db 1219 TAGGCTGCTATAGTATGATGATCATCTGCGGCTTTAGCTTGTCTCTTATGTTTC 1278  
Qy 904 TTTAATCAGATTAACCTGCGGAGCCTTTTCAATTTAAGCGCTGGAAGCAGTCTTTTGG 963  
Db 1279 TTTAATCAGATTAACCTGCGGAGCCTTTTCAATTTAAGCGCTGGAAGCAGTCTTTTGG 1338  
Qy 964 CTAGTTAATTAAGTGTGTTTTCGTTAATTAAGCAAAATTAATTAATTAATTAATG 1020  
Db 1339 CTAGTTAATTAAGTGTGTTTTCGTTAATTAAGCAAAATTAATTAATTAATTAATG 1395

RESULT 5  
US-10-948-518-23  
Sequence 23, Application US/10948518  
Publication No. US20050064492A1  
GENERAL INFORMATION:  
APPLICANT: FREDERIC J. DESAUVAGE  
APPLICANT: GRETCHEN FRANTZ  
APPLICANT: KENNETH J. HILLAN  
APPLICANT: PAUL POLAKIS  
APPLICANT: ANDREW POLSON  
APPLICANT: VICTORIA SMITH  
APPLICANT: SUSAN D. SPENCER  
APPLICANT: THOMAS D. WU  
APPLICANT: ZEMIN ZHANG  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND  
TREATMENT OF TUMOR  
FILE REFERENCE: P5026R1-US  
CURRENT APPLICATION NUMBER: US/10/948,518  
CURRENT FILING DATE: 2004-09-22  
PRIORITY APPLICATION NUMBER: US/10/643,795  
PRIORITY FILING DATE: 2003-08-19  
PRIORITY APPLICATION NUMBER: US 60/404,809  
PRIORITY FILING DATE: 2002-08-19  
PRIORITY APPLICATION NUMBER: US 60/405,645  
PRIORITY FILING DATE: 2002-08-21  
PRIORITY APPLICATION NUMBER: US 60/413,192  
PRIORITY FILING DATE: 2002-09-23  
PRIORITY APPLICATION NUMBER: US 60/419,008  
PRIORITY FILING DATE: 2002-10-15  
PRIORITY APPLICATION NUMBER: US 60/426,847  
PRIORITY FILING DATE: 2002-11-15  
PRIORITY APPLICATION NUMBER: US 60/484,959  
PRIORITY FILING DATE: 2003-07-02  
NUMBER OF SEQ ID NOS: 158  
SEQ ID NO 23  
LENGTH: 1395  
TYPE: DNA  
ORGANISM: Homo sapien  
US-10-948-518-23

Query Match 98.7%; Score 1013.8; DB 9; Length 1395;  
Best Local Similarity 99.8%; Pred. No. 1.1e-289;  
Matches 1015; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 4 CAAAGTTGGGCAAAAAATCAAGTATTTGGTCTCCGAAACAAAGCTTATCATACAT 63
DB 379 CAAAGTTGGGCAAAAAATCAAGTATTTGGTCTCCGAAACAAAGCTTATCATACAT 438
QY 64 AAACAATTGATGAGATGTTTCCCAAGCCCACTATTTTCTTCTTCAATTGGCGAA 123
DB 439 AAACAATTGATGAGATGTTTCCCAAGCCCACTATTTTCTTCTTCAATTGGCGAA 498
QY 124 ACAAGCTCAGAGGCTGGAACATACCTTGTCTTCTTGAATAATTTTCCCTGATGT 183
DB 499 ACAAGCTCAGAGGCTGGAACATACCTTGTCTTCTTGAATAATTTTCCCTGATGT 558
QY 184 ATTAAATACATTGGCAAGAAAGAGCAACGATTTGGGATCCAGAGGGGAAAC 243
DB 559 ATTAAATACATTGGCAAGAAAGAGCAACGATTTGGGATCCAGAGGGGAAAC 618
QY 244 ACCATGAAGCTAACGACATACATGAAATTTAGCTGTAAACGGTCCAGAAAGTCA 303
DB 619 ACCATGAAGCTAACGACATACATGAAATTTAGCTGTAAACGGTCCAGAAAGTCA 678
QY 304 CTGGAACAAGACACAGATGATCGTCAACATGAGATTAATAAAACGAGTTGATCAA 363
DB 679 CTGGAACAAGACACAGATGATCGTCAACATGAGATTAATAAAACGAGTTGATCAA 738
QY 364 GAAATTAATCTTCTCCATTAAGACGAGTGTCAATGATGCCAAAGACATTTGT 423
DB 739 GAAATTAATCTTCTCCATTAAGACAGATGTCAATGATGCCAAAGACATTTGT 798
QY 424 TCAAAAAGTCAATGATGATCACTCTGCTGAGTCAAAAACCTCTGATTAATG 483
DB 799 TCAAAAAGTCAATGATGATCACTCTGCTGAGTCAAAAACCTCTGATTAATG 858
QY 484 TACCTCTCTGCTCTCAAGAGTGTGATTAATTTCCATCATCATCTGCTGCTCT 543
DB 859 TACCTCTCTGCTCTCAAGAGTGTGATTAATTTCCATCATCATCTGCTGCTCT 918
QY 544 AGAAGAACGGCTTTCTGCTGATGAGAGAAATCATTAACAGCGGTGGCAAGAGAGC 603
DB 919 AGAAGAACGGCTTTCTGCTGATGAGAGAAATCATTAACAGCGGTGGCAAGAGAGC 978
QY 604 CATCTTTTCTCATCGGTATTTGCTTGAAGAGCTCTTCTGAGAGATCTAGTGGCTTT 663
DB 979 CATCTTTTCTCATCGGTATTTGCTTGAAGAGCTCTTCTGAGAGATCTAGTGGCTTT 1038
QY 664 CTTTCTGAGGTTTGGGCAATTCAGTTCATGATGATTAATTTCTATCATTAATGTAATA 723
DB 1039 CTTTCTGAGGTTTGGGCAATTCAGTTCATGATGATTAATTTCTATCATTAATGTAATA 1098
QY 724 CGGTTTTCAACCAAGTGGGCAACAGAGAACTCTCATCTGTAAATTAAGATGAGATAGC 783
DB 1099 CGGTTTTCAACCAAGTGGGCAACAGAGAACTCTCATCTGTAAATTAAGATGAGATAGC 1158
QY 784 CAGGCGATCTCCAGACCAATCTCTCATGTTTTTCCAGAGCTCTTCAAGCAACCCAAA 843
DB 1159 CAGGCGATCTCCAGACCAATCTCTCATGTTTTTCCAGAGCTCTTCAAGCAACCCAAA 1218
QY 844 TAGGCGCTGATAGTGAACATCTGCGGCTTCTAGGCTTGTCCCTCTTATGATGTTTC 903
DB 1219 TAGGCGCTGATAGTGAACATCTGCGGCTTCTAGGCTTGTCCCTCTTATGATGTTTC 1278
QY 904 TTTAATCAGATTAATGCTGAGAACCTTTCAATTTTAAACGCGCTGAGAGAGTCTTTCTTTG 963
DB 1279 TTTAATCAGATTAATGCTGAGAACCTTTCAATTTTAAACGCGCTGAGAGAGTCTTTCTTTG 1338
QY 964 CTAAGTGAATTAATGCTGCTTTTCCGTAAATAAGCAAAATTAATTTAAAAAATG 1020
DB 1339 CTAAGTGAATTAATGCTGCTTTTCCGTAAATAAGCAAAATTAATTTAAAAAATG 1395
```

RESULT 6  
US-10-101-510-492  
Sequence 492, Application US/10101510  
Publication No. US20030148295A1  
GENERAL INFORMATION:  
APPLICANT: WAN, JACKSON  
APPLICANT: WANG, YIXIN  
TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE  
FILE REFERENCE: 15117.0012  
CURRENT FILING DATE: 2002-03-20  
PRIOR FILING DATE: 2001-03-20  
NUMBER OF SEQ ID NOS: 805  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 492  
LENGTH: 1799  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-101-510-492

Query Match 94.3%; Score 968.2; DB 6; Length 1799;  
Best Local Similarity 98.1%; Pred. No. 4.3e-276;  
Matches 990; Conservative 0; Mismatches 18; Indels 1; Gaps 1;

```
QY 20 AAATCAAGTATTTGGTCCGGAACAAGCTTATCATTAAGATTAACAACCTGATGACAG 79
DB 791 ATAAGAACTCTTTGGCAGTGAACAACACTTGTGTACAGATTAACAACCTGATGACAG 850
QY 80 ATGTTTCCCAAGCCCACTATTTTCTTCTTCAATGCTGGAACAAAGCTCAGAAAG 139
DB 851 ATGTTTCCCAAGCCCACTATTTTCTTCTTCAATGCTGGAACAAAGCTCAGAAAG 910
QY 140 CTGGAACATACCTTTTGTCTTCTTGAAGAAATTTTCCCTGATGTTAATTAATCATTTGGC 199
DB 911 CTGGAACATACCTTTTGTCTTCTTGAAGAAATTTTCCCTGATGTTAATTAATCATTTGGC 970
QY 200 AAGAAAAGAGCAACAGATTTCT-GGATCCAGAGAGGGAACAACATGAAGACTAAC 258
DB 971 AAGAAAAGAGCAACAGATTTCTGGGATCCAGAGAGGGAACAACATGAAGACTAAC 1030
QY 259 GACACATACATGAATTTTATGCTGTTAACGCTCCAGAAAAGTCACTGACAAAGAAC 318
DB 1031 GACACATACATGAATTTTATGCTGTTAACGCTCCAGAAAAGTCACTGACAAAGAAC 1090
QY 319 AGATGATCTGACACATGAAATTAATAAAGGAGTTGATGAAGAAATTAATCTTCTCT 378
DB 1091 AGATGATCTGACACATGAAATTAATAAAGGAGTTGATGAAGAAATTAATCTTCTCT 1150
QY 379 CCAATTAAGCGGATGTCATCAATGATCCCAAGAACAAATGTTCAAAAGATGCAAT 438
DB 1151 CCAATTAAGCGGATGTCATCAATGATCCCAAGAACAAATGTTCAAAAGATGCAAT 1210
QY 439 GATACACTACTGCTGAGCTCACAAAACCTCTGCAATTAATTAATCACTCTCTGCTC 498
DB 1211 GATACACTACTGCTGAGCTCACAAAACCTCTGCAATTAATTAATCACTCTCTGCTC 1270
QY 499 CTCAGAGTGTGATTAATTTTGCATCATCACTGCTGTCTGTTAGAAAGAGGCTTTTC 558
DB 1271 CTCAGAGTGTGATTAATTTTGCATCATCACTGCTGTCTGTTAGAAAGAGGCTTTTC 1330
QY 559 TGTGCAATGGAAGAAATCATTAACAGAGGTTGGCAACAGAGAGCCATCTTTCTCATC 618
DB 1331 TGTGCAATGGAAGAAATCATTAACAGAGGTTGGCAACAGAGAGCCATCTTTCTCATC 1390
QY 619 GGTATTTGCTCTAGAGAGCTCTTCTGAGGATCTAGTTGGGCTTTCTTCTGAGTTTGGG 678
DB 1391 GGTATTTGCTCTAGAGAGCTCTTCTGAGGATCTAGTTGGGCTTTCTTCTGAGTTTGGG 1450
QY 679 CCAATTCAGTTCAATGCTGATCAATTAATTAATTAATTAATTAATTAATTAATTAAT 738
DB 1451 CCAATTCAGTTCAATGCTGATCAATTAATTAATTAATTAATTAATTAATTAATTAAT 1510
```

Qy 739 TGGGACACAGAGAACTCTCTGTATTAACAATGAGGAATAGCCACGCGGATCTCCAG 798  
Db 1511 TGGGACACAGAGAACTCTCTGTATTAACAATGAGGAATAGCCACGCGGATCTCCAG 1570  
Qy 799 CACCAATCTCTCCATGTTTTCCACAGCTCCCTCCAGCCCAACCAATAGCGCTGTATG 858  
Db 1571 CACCAATCTCTCCATGTTTTCCACAGCTCCCTCCAGCCCAACCAATAGCGCTGTATG 1630  
Qy 859 TGTAGACATCTCTGCGGCTTCTAGCCTTGTCCCTCTCTTATGTTCTTTAATCAGATACT 918  
Db 1531 TGTAGACATCTCTGCGGCTTCTAGCCTTGTCCCTCTCTTATGTTCTTTAATCAGATACT 1690  
Qy 919 GCCTGGAAGCCTTCTTATTTTACAGCCCTGGAAGCATCTTCTTGCTAGTGAATATGT 978  
Db 1591 GCCTGGAAGCCTTCTTATTTTACAGCCCTGGAAGCATCTTCTTGCTAGTGAATATGT 1750  
Qy 979 GGTGTGTTTTCCGTAATAGCAAAATTAATTTAAAAAATGAAAAGTT 1027  
Db 1751 GGTGTGTTTTCCGTAATAGCAAAATTAATTTAAAAAATGAAAAGTT 1799

## RESULT 7

US-09-957-708-16  
; Sequence 16, Application US/09957708  
; Publication No. US20030031678A1  
; GENERAL INFORMATION:  
; APPLICANT: Sun, Yongming  
; APPLICANT: Recipon, Hervé  
; APPLICANT: Cafferey, Robert  
; APPLICANT: Ali, Shujath  
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific  
; TITLE OF INVENTION: Genes  
; FILE REFERENCE: DEX-0239  
; CURRENT APPLICATION NUMBER: US/09/957,708  
; CURRENT FILING DATE: 2001-09-19  
; PRIOR APPLICATION NUMBER: 60/233,746  
; PRIOR FILING DATE: 2000-09-19  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 16  
; LENGTH: 1418  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-957-708-16

Query Match 92.1%; Score 945.6; DB 3; Length 1418;  
Best Local Similarity 97.8%; Pred. No. 1,9e-269;  
Matches 1001; Conservative 0; Mismatches 19; Indels 4; Gaps 4;

Qy 4 CAAGAGTTGGCCAAAAAATCAAGGTATTTGGTCCGGACAAAGCTTATCATTAACAGAT 63  
Db 379 CAAGAGTTGGCCAAAAAATCAAGGTATTTGGTCCGGACAAAGCTTATCATTAACAGAT 438  
Qy 64 AAAACACTTGATGAGATGTTTCCCGCAAGCCATATTTTCTTCCCTTCAATTCGTA 123  
Db 439 AAAACACTTGATGAGATGTTTCCCGCAAGCCATATTTTCTTCCCTTCAATTCGTA 498  
Qy 124 ACAAGCTCCAGAAAGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGT 183  
Db 499 ACAAGCTCCAGAAAGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGT 558  
Qy 184 ATTAAGATACATTTGGCAAGAAAAAGAGCAACAGATTTGGGATCCAGAGGGGAAC 243  
Db 559 ATTAAGATACATTTGGCAAGAAAAAGAGCAACAGATTTGGGATCCAGAGGGGAAC 618  
Qy 244 ACCATGAAGACTTAAGACACATACATGAATTTAGCTGTGAACGCTCCGAAAAAGTCA 303  
Db 619 ACCATGAAGACTTAAGACACATACATGAATTTAGCTGTGAACGCTCCGAAAAAGTCA 678  
Qy 304 CTGGAACAAGAACACAGATGTATGCTCAGACATGAGAAATTAATAAAAACGAGTTGATCA 363  
Db 679 CTGGAACAAGAACACAGATGTATGCTCAGACATGAGAAATTAATAAAAACGAGTTGATCA 738

Qy 364 GAAATTTATCTTTCTCTCCATTAAGACGATGTCTATCATATGATTCCTCAAGACATTTGT 423  
Db 739 GAAATTTATCTTTCTCTCCATTAAGACGATGTCTATCATATGATTCCTCAAGACATTTGT 798  
Qy 424 TCAAAAGATCAATATGATTAATCACTACTGTGAGCTCAAAACACTCTGCATATTAATG 483  
Db 799 TCAAAAGATCAATATGATTAATCACTACTGTGAGCTCAAAACACTCTGCATATTAATG 858  
Qy 484 TACCTCTCTGCTCTCTCAAGAGTGTGTATTTTGGCATCATACCTGTCTGTCTT 543  
Db 859 TACCTCTCTGCTCTCTCAAGAGTGTGTATTTTGGCATCATACCTGTCTGTCTT 918  
Qy 544 AGAAGAACGCTTTCTCTGCAATGAGAGAAATCATTAACAGAGGTGGACAAAGAGGC 603  
Db 919 AGAAGAACGCTTTCTCTGCTCAATGAGAGAAATCATTAACAGAGGTGGACAAAGAGGC 978  
Qy 604 CATCTTTCTCTC-ATCGGATATTGTCCTTGAAGCGCTTCTGAGATCTAGTTGGCTT 662  
Db 979 CATCTTTCTCTCAATCGGATATTGTCCTTGAAGCGCTTCTGAGATCTAGTTGGCTT 1038  
Qy 663 TCTTCTGGGTTTGGGCAATTCAGTTCTCATGTGTACTATT-CTATCATTAATGTA- 720  
Db 1039 TCTTCTGGGTTTGGGCAATTCAGTTCTCATGTGTACTATTCTATCATTAATGTA- 1098  
Qy 721 TAACGTTTTCAAACCAAGTGGGACACAGAGAACTCATCTGTGAATTAACAATGAGAA 780  
Db 1099 TAACGTTTTCAAACCAAGTGGGACACAGAGAACTCATCTGTGAATTAACAATGAGAA 1158  
Qy 781 AGCCACGGGATCTCCAGACCAATCTCTCATGTTTTCACAGAGCTCTCCAGCAACC 840  
Db 1159 AGCCACGGGATCTCCAGACCAATCTCTCATGTTTTCACAGAGCTCTCCAGCAACC 1218  
Qy 841 AAATAGGCTGCTATATAGTGTAGACATCTGCGGCTTCTAGCCTTGTCTCTTAAGT 900  
Db 1219 AAATAGGCTGCTATATAGTGTAGACATCTGCGGCTTCTAGCCTTGTCTCTTAAGT 1277  
Qy 901 TTTCTTATCAGATACATGCTCTGGAAGCTTTTCAATTTTACAGCCCTGGAAGCATCTTCT 960  
Db 1278 TTTCTTATCAGATACATGCTCTGGAAGCTTTTCAATTTTACAGCCCTGGAAGCATCTTCT 1337  
Qy 961 TTGCTAGTTGAATATGATGATGTTTTCGTAATAGCAAAATTAATTTAAAAAATG 1020  
Db 1338 TTGCTAGTTGAATATGATGATGTTTTCGTAATAGCAAAATTAATTTAAAAAAT 1397  
Qy 1021 AAAA 1024  
Db 1398 GAAA 1401

## RESULT 8

US-09-925-300-58  
; Sequence 58, Application US/09925300  
; Patent No. US20020151681A1  
; GENERAL INFORMATION:  
; APPLICANT: Craig Ruben,  
; APPLICANT: Steve Rosen,  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA101  
; CURRENT APPLICATION NUMBER: US/09/925,300  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05988  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1890  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 58  
; LENGTH: 1155  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (135)

OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: misc feature  
LOCATION: (432)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: misc feature  
LOCATION: (443)  
OTHER INFORMATION: n equals a,t,g, or c  
US-09-925-300-58

Query Match 88.2%; Score 906; DB 3; Length 1155;  
Best Local Similarity 97.8%; Pred. No. 9.8e-258;  
Matches 951; Conservative 8; Mismatches 8; Indels 5; Gaps 4;

59 CAGATAACCACTGATGAGATGTTCCCAAGCCACATATTTCTTCCCTGATG 118  
160 CTGTAACCACTGATGAGATGTTCCCAAGCCACATATTTCTTCCCTGATG 218  
119 CTGTAACCACTGATGAGATGTTCCCAAGCCACATATTTCTTCCCTGATG 178  
219 CTGTAACCACTGATGAGATGTTCCCAAGCCACATATTTCTTCCCTGATG 278  
179 ATGTTATTAAGATGATGAGATGTTCCCAAGCCACATATTTCTTCCCTGATG 238  
279 ATGTTATTAAGATGATGAGATGTTCCCAAGCCACATATTTCTTCCCTGATG 338  
239 GG--AAACCACTGATGAGATGTTCCCAAGCCACATATTTCTTCCCTGATG 296  
339 GGGAACCACTGATGAGATGTTCCCAAGCCACATATTTCTTCCCTGATG 398  
297 AAATGCTGATGAGATGTTCCCAAGCCACATATTTCTTCCCTGATG 354  
399 AAATGCTGATGAGATGTTCCCAAGCCACATATTTCTTCCCTGATG 458  
355 GTTGTATGAGATGTTCCCTGATGAGATGTTCCCAAGCCACATATTTCTTCCCTGATG 414  
459 GTTGTATGAGATGTTCCCTGATGAGATGTTCCCAAGCCACATATTTCTTCCCTGATG 518  
415 GACAAATGTTCAAAAGATGAGATGTTCCCTGATGAGATGTTCCCAAGCCACATATTTCTTCCCTGATG 474  
519 GACAAATGTTCAAAAGATGAGATGTTCCCTGATGAGATGTTCCCAAGCCACATATTTCTTCCCTGATG 578  
475 TATTACATGATGATGTTCCCTGATGAGATGTTCCCTGATGAGATGTTCCCTGATG 534  
579 TATTACATGATGATGTTCCCTGATGAGATGTTCCCTGATGAGATGTTCCCTGATG 638  
535 TGTCTGTTAGAGAGATGTTCCCTGATGAGATGTTCCCTGATGAGATGTTCCCTGATG 594  
639 TGTCTGTTAGAGAGATGTTCCCTGATGAGATGTTCCCTGATGAGATGTTCCCTGATG 698  
595 CAAGAGAGCCATCTTTTCTGATGAGATGTTCCCTGATGAGATGTTCCCTGATG 654  
699 CAAGAGAGCCATCTTTTCTGATGAGATGTTCCCTGATGAGATGTTCCCTGATG 758  
655 TTGGGCTTTCTTCTGATGAGATGTTCCCTGATGAGATGTTCCCTGATGAGATGTTCCCTGATG 714  
759 TTGGGCTTTCTTCTGATGAGATGTTCCCTGATGAGATGTTCCCTGATGAGATGTTCCCTGATG 818  
715 ATTGTATTAAGATGTTCCCTGATGAGATGTTCCCTGATGAGATGTTCCCTGATG 774  
819 ATTGTATTAAGATGTTCCCTGATGAGATGTTCCCTGATGAGATGTTCCCTGATG 878  
775 AGGAATAGCCAGATGTTCCCTGATGAGATGTTCCCTGATGAGATGTTCCCTGATG 834  
879 AGGAATAGCCAGATGTTCCCTGATGAGATGTTCCCTGATGAGATGTTCCCTGATG 938  
835 CAACCAATAGCCAGATGTTCCCTGATGAGATGTTCCCTGATGAGATGTTCCCTGATG 894  
939 CAACCAATAGCCAGATGTTCCCTGATGAGATGTTCCCTGATGAGATGTTCCCTGATG 998  
895 TTAGTGTCTTTTAAATGATGAGATGTTCCCTGATGAGATGTTCCCTGATGAGATGTTCCCTGATG 954  
999 TTAGTGTCTTTTAAATGATGAGATGTTCCCTGATGAGATGTTCCCTGATGAGATGTTCCCTGATG 1058

955 TCTTCTTTGCTAGTGAATTAATGATGTTTTCCTGTAATAGCAAAATTAATTTAA 1014  
1059 TCTTCTTTGCTAGTGAATTAATGATGTTTTCCTGTAATAGCAAAATTAATTTAA 1118  
1015 AAATGAAAAGT 1026  
1119 AAATGAAAARW 1130

## RESULT 9

US-09-954-456-317  
Sequence 317, Application US/09954456  
Patent No. US20020115057A1  
GENERAL INFORMATION:  
APPLICANT: Young, Paul  
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc  
FILE REFERENCE: 689290-76  
CURRENT FILING DATE: 2001-09-18  
PRIOR FILING DATE: US/60/233,617  
PRIOR FILING DATE: 2000-09-18  
PRIOR APPLICATION NUMBER: US/60/234,052  
PRIOR FILING DATE: 2000-09-20  
PRIOR APPLICATION NUMBER: US/60/234,923  
PRIOR FILING DATE: 2000-09-25  
PRIOR APPLICATION NUMBER: US/60/235,134  
PRIOR FILING DATE: 2000-09-25  
PRIOR APPLICATION NUMBER: US/60/235,637  
PRIOR FILING DATE: 2000-09-26  
PRIOR APPLICATION NUMBER: US/60/235,638  
PRIOR FILING DATE: 2000-09-26  
PRIOR APPLICATION NUMBER: US/60/235,711  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: US/60/235,720  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: US/60/235,840  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: US/60/235,863  
PRIOR FILING DATE: 2000-09-27  
NUMBER OF SEQ ID NOS: 2276  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 317  
LENGTH: 1421  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-954-456-317

Query Match 87.8%; Score 902.2; DB 3; Length 1421;  
Best Local Similarity 92.8%; Pred. No. 1.5e-256;  
Matches 977; Conservative 0; Mismatches 28; Indels 48; Gaps 1;

20 AAATCAAGTATTTGTCCTGGAACCAAGCTTATCAATTAAGATTAACCACTGATG 79  
369 ATAAGAACTCTTTGGAGTGAACAACATTTGTCACATTAACCACTGATG 428  
80 ATGTTTCCCAAGCCATATTTTCTTCTTCAATGCTGAACCAAGCTGCAAGG 139  
429 ATGTTTCCCAAGCCATATTTTCTTCTTCAATGCTGAACCAAGCTGCAAGG 488  
140 CTGGAACATACCTTTGCTTCTGAGAAATTTTCCCTGATGTTATTAAGATGATG 199  
489 CTGGAACATACCTTTGCTTCTGAGAAATTTTCCCTGATGTTATTAAGATGATG 548  
200 AAGAAAAGAGCAACAGATTTCTGGATCCAGAGGGGAAACCATGAAGCTAAG 259  
549 AAGAAAAGAGCAACAGATTTCTGGATCCAGAGGGGAAACCATGAAGCTAAG 608  
260 ACACATACATGAATTTAGTGTGTTAAGCTGCAAGAAAGTCACTGAGCAAAAG 319  
609 ACACATACATGAATTTAGTGTGTTAAGCTGCAAGAAAGTCACTGAGCAAAAG 668  
320 GATGTATGTCAGATGATGAATTAATTAAGAAAGGATGATCAAGAAATATCTTCTC 379



Db	1089	CATGTGTGTA	CTATTCTATCA	TTTATGTATAT	ATGSGTTTTT	CAAAACAG	TGGGGCACA	GAG	1148	
Qy	752	AACTTCACT	CTGTATAT	TAACAAT	TAGGAAT	TAGCCAG	CGGAGAT	CTCCAGCA	CAATCTCTCC	811
Db	1149	AACTCAGT	CTGTATAT	TAACAAT	TAGGAAT	TAGCCAG	TGGGATCT	CCAGACCA	CAATCTCTCC	1208
Qy	812	ATGTTTTCA	CAGCTCT	CCAGCCA	ACC	CAATAT	AGCGCTCT	ATATGTA	GAGATCCTG	871
Db	1209	ATGTTTTCA	CAGCTCT	CCAGCCA	ACC	CAATAT	AGCGCTCT	ATATGTA	GAGACGCTG	1268
Qy	872	CGGTTTCA	GCGTTGT	CCCTCT	CTTAA	GTCTT	TAATCA	ATACTG	CCCTGGAA	931
Db	1269	CGGTTTCA	GCGTTGT	CCCTCT	CTTAA	GTCTT	TAATCA	ATACTG	CCCTGGAA	1328
Qy	932	TCATTTTACA	CGCCCTGA	AGCACTT	CTTGT	CTAGT	AGTAAT	TTATGT	GTGTTTTTCC	991
Db	1329	TCATTTTACA	CGCCCTGA	AGCACTT	CTTGT	CTAGT	AGTAAT	TTATGT	GTGTTTTTCC	1388
Qy	992	GTAATTA	GCAAAAT	TAATTTT	AAAAA	AGAAAA				1024
Db	1389	GTAATTA	GCAAAAT	TAATTTT	AAAAA	AGAAAA				1421

```

RESULT 11
US-09-960-706-676
; Sequence 676, Application US/09960706
; Publication No. US20030134280A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasia
; TITLE OF INVENTION: Gene Expression Profiles
; FILE REFERENCE: 44921-5029-01US
; CURRENT APPLICATION NUMBER: US/09/960, 706
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 60/223,323
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 09/873,319
; PRIOR FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 676
; LENGTH: 1586
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134280A1 M30894
; US-09-960-706-676

```

	Query Match	81.6%	Score 838	DB 3	Length 1586
	Best Local Similarity	88.4%	Pred. No.1.8e-237		
	Matches	962	Conservative	0	Mismatches 30; Indels 96; Gaps 14
Qy	20	AAATCAAGGATTTATTTGGTCCCGAAACAAGCTTATCATACAGATTAACAACCTTGATGAG	79		
Db	499	ATAAAGAACTCTTTGGCAGTGGAAACAACCTTTGTGACAGATTAACAACCTTGATGAG	558		
Qy	80	ATGTTTCCCGCAAGCCCACTATTTTCTTCCTTCAATTCTGAAACAAGCTCCGAAG	139		
Db	559	ATGTTTCCCGCAAGCCCACTATTTTCTTCCTTGATTTCTGAAACAAGCTCCGAAG	618		
Qy	140	CTGGAACTACCTTTGTTCTTCTTGAAATTTTCCCTGAGTTTATTAGATCAATTGGC	199		
Db	619	CTGGAACTATCTTTGTTCTTCTTGAAATTTTCCCAAGATTATTATTAGATCAATTGGC	678		
Qy	200	AAGAAGAAAGCAACAGATTCCTGGGATCCCAAGAGGGGAACCACTGAAGACTAAG	259		
Db	679	AAGAAGAAAGCAACAGATTCCTGGGATCCCAAGAGGGGAACCACTGAAGACTAAG	738		
Qy	260	ACACATACATGAATTTTGTGCTGTTAAAGGTGCCAGAAAGCTCATGTGACAAAGAACA	319		
Db	739	ACACATACATGAATTTTGTGCTGTTAAAGGTGCCAGAAAGCTCATGTGACAAAGAACA	798		
Qy	320	GATGATTCGTGACATGAGATTAATAAAAAAGAGTTGATCAAGAAATTATCTTTCCTC	379		

Db	799	GATGATTCGTCAGACATGAGAAATTAATAAAAAAGAAATTTGATTCAGAAATTAATCTTCTC	858
Qy	380	CAATPAAAGAC-----	389
Db	859	CAATPAAAGACAGATGTCACCCACAGTGTATCCAAAGACAGTTATTCAAAAGATGCAATG	918
Qy	390	-----	403
Db	919	ATGTACACACAGTGTATCCAAATACAAATTAATTCMAAGATGCAAAATGATGTCTACAA	978
Qy	404	TGATATCCAAAGACAAATTTGTTCAAAAGATGCMAATGATACACTACTCTGACGCTCAAA	463
Db	979	TGGATCCCAAGACAAATTTGGTCAAAAGATGCMAATGATACACTACTCTGACGCTCAAA	1038
Qy	464	ACACCTCTGCATATTACATGTACTCTCTCTGCTCTCAAGATGTGATCTAATTTTGCA	523
Db	1039	ACACTCTGCATATTACATGTACTCTCTCTGCTCTCAAGATGTGATCTAATTTTGCA	1098
Qy	524	TCATACCTCTGCTCTCTGCTTGAAGAAACGGCTTTCTGCTGCATATGAGAGAAATCATATAC	583
Db	1099	TCATACCTCTGCTCTCTGCTTGAAGAAACGGCTTTCTGCTGCATATGAGAGAAATCATATAC	1158
Qy	584	AGAGCGTGGCAACAGAGAGGCCATCTTTTCCATCGGTATATGTCCCTAGAGGCTCTC	643
Db	1159	AGAGCGTGGCAACAGAGAGGCCATCTTTTCCATCGGTATATGTCCCTAGAGGCTCTC	1218
Qy	644	TGAGGATCTAGATGGGCTTTCTTTCTGCGTTTGGGCAATTCAGTTCTCATGTGTACT	703
Db	1219	TGAGGATCTAGATGGGCTTTCTTTCTGCGTTTGGGCAATTCAGTTCTCATGTGTACT	1278
Qy	704	ATTCTATCATTTATTTGTATTAACGGTTTTCAACACAGTGGGCAACAGAGAACTCATCTG	763
Db	1279	ATTCTATCATTTATTTGTATTAACGGTTTTCAACACAGTGGGCAACAGAGAACTCATCTG	1338
Qy	764	TAAATPAAATAGAGAAATAGCCACGGCGATCTCCAGACCAATCTCTCAGTTTTTCCACA	823
Db	1339	TAAATPAAATAGAGAAATAGCCATGCGCATCTCCAGACCAATCTCTCAGTTTTTCCACA	1398
Qy	824	GCTCCTCAGCAACCCAAATATAGCGCTGCTATATGATATGACATCTGCGGCTCTAGCC	883
Db	1399	GCTCCTCAGCAACCCAAATATAGCGCTGCTATATGATATGACATCTGCGGCTCTAGCC	1458
Qy	884	TTGTCCCTCTTATGTTCTTTTATCAGATPAACTGCGTGAAGCCTTTCAATTTACAG	943
Db	1459	TTGTCCCTCTTATGTTCTTTTATCAGATPAACTGCGTGAAGCCTTTCAATTTACAG	1518
Qy	944	CCCTGAAGCACTCTTTCTTGTACTGTAATATGTGTGTCTTTTCCGTAATAGCAA	1003
Db	1519	CCCTGAAGCACTCTTTCTTGTACTGTAATATGTGTGTCTTTTCCGTAATAGCAA	1578
Qy	1004	ATAAATTT 1011	
Db	1579	ATAAATTT 1586	

```

RESULT 12
US-09-873-319-424
; Sequence 424, Application US/09873319A
; Publication No. US20030134324v1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; APPLICANT: Kulкарт, Prakash
; APPLICANT: Getzenberg, Robert H.
; APPLICANT: Waga, Iwao
; APPLICANT: Yamamoto, Jun
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic
; TITLE OF INVENTION: Hyperplasia Using Gene Expression Profiles
; FILE REFERENCE: 44921-5029-US
; CURRENT APPLICATION NUMBER: US/09/873.319A
; CURRENT FILING DATE: 2001-06-05
; EARLIER APPLICATION NUMBER: US 60/223,323
; EARLIER FILING DATE: 2000-08-07

```

NUMBER OF SEQ ID NOS: 755  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO: 424  
LENGTH: 1586  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: Genbank Accession No. US2003013432A1 M30894  
US-09-873-319-424

Query Match 81.6%; Score 838; DB 3; Length 1586;  
Best Local Similarity 88.4%; Pred. No. 1.8e-237;  
Matches 962; Conservative 0; Mismatches 30; Indels 96; Gaps 1;

```
QY 20 AAATCAAGTATTTGTCCTCCGGAACAAAGCTTATCATACAGATTAACAACCTTGATGACG 79
DB 499 ATAAAGAACTCTTTGGCAGTGAACAACCTTGTTGTACAGATTAACAACCTTGATGACG 558
QY 80 ATGTTTCCCAAGCCCACTATTTTCTTCTCAATTGCTGAACAAAGCTCCAGAGG 139
DB 559 ATGTTTCCCAAGCCCACTATTTTCTTCTCAATTGCTGAACAAAGCTCCAGAGG 618
QY 140 CTGAACATACCTTTGCTTCTGAGAAATTTTCCCGATGTTATTAAGATGATGACG 199
DB 619 CTGAACATATCTTTGCTTCTGAGAAATTTTCCCGATGTTATTAAGATGATGACG 678
QY 200 AAGAAAGAGAGACACAGATTCGGATCCAGAGGGAGACCATGAAGACTAAGC 259
DB 679 AAGAAAGAGAGACACAGATTCGGATCCAGAGGGAGACCATGAAGACTAAGC 738
QY 260 ACACATACATGAATTTTACCTGCTTACCGTGCAGAAAGTCACTGACAAAGACACA 319
DB 739 ACACATACATGAATTTTACCTGCTTACCGTGCAGAAAGTCACTGACAAAGACACA 798
QY 320 GATGATCGTGAACATGAGAAATTAATAACGGAGTTGATCAAGAAATTTCTTCTC 379
DB 799 GATGATCGTGAACATGAGAAATTAATAACGGAGTTGATCAAGAAATTTCTTCTC 858
QY 380 CAATTAAGAC----- 389
DB 859 CAATTAAGACAGATGTACACAGATGATCCCAAGACAGTTATTCAAAAGATGCAATG 918
QY 390 -----GGATGTCATCAAA 403
DB 919 ATGTCAACACAGATGATCCCAATTAATTAACAAGATGCAAAATGATGTCATCAAA 978
QY 404 TGGATCCCAAGACAAATTTGTTCAAAAGATGAATATACCTACTGTGAGAGCTCAAA 463
DB 979 TGGATCCCAAGACAAATTTGTTCAAAAGATGAATATACCTACTGTGAGAGCTCAAA 1038
QY 464 ACACCTTGCAATATTAATGATGATCTCTCTGCTCTCAAGAGTGTGTTATTGCGCA 523
DB 1039 ACACCTTGCAATATTAATGATGATCTCTCTGCTCTCAAGAGTGTGTTATTGCGCA 1098
QY 524 TCATCACTGCTGTCTGTTAGAGAGCGGCTTTCTGCTGCAATGAGAGAAATCATAC 583
DB 1099 TCATCACTGCTGTCTGTTAGAGAGCGGCTTTCTGCTGCAATGAGAGAAATCATAC 1158
QY 584 AGACGCTGAGCAACAAGAGCGCATCTTTCTCATGCGTTATTTGCCNAGAGCTCTC 643
DB 1159 AGACGCTGAGCAACAAGAGCGCATCTTTCTCATGCGTTATTTGCCNAGAGCTCTC 1218
QY 644 TGAAGATCAAGTGGGCTTTCTTTCTGGGTTTGGGCAATTCAGTTCTCATGTGTACT 703
DB 1219 TGAAGATCAAGTGGGCTTTCTTTCTGGGTTTGGGCAATTCAGTTCTCATGTGTACT 1278
QY 704 ATTCTATCATTAATTTGATTAACGGTTTCAACACAGTGGGCAACAAGAGACCTCACTG 763
DB 1279 ATTCTATCATTAATTTGATTAATGTTTCAACACAGTGGGCAACAAGAGACCTCACTG 1338
QY 764 TAAATTAACAATGAGAAATAGCCAGCGGATCTCCAGACCAATCTCTCATGTTTCCACA 823
DB 1339 TAAATTAACAATGAGAAATAGCCAGCGGATCTCCAGACCAATCTCTCATGTTTCCACA 1398
```

```
QY 824 GCTCTTCACGCAACCCAAATAGCGCTGCTATAGTGTAGACATCCCTGGGCTTTAGCC 883
DB 1399 GCTCTTCACGCAACCCAAATAGCGCTGCTATAGTGTAGACAGCCCTGGGCTTTAGCC 1458
QY 884 TTGTCCCTCTTATAGTGTCTTTTATTCAGATTAACCTGCTGGAAGCCCTTCAATTTACG 943
DB 1459 TTGTCCCTCTTATAGTGTCTTTTATTCAGATTAACCTGCTGGAAGCCCTTCAATTTACG 1518
QY 944 CCTGAAGCAGTCTTCTTCTGCTAGTTGAATTAATGATGATGTTTCCGTAATTAAGCAA 1003
DB 1519 CCTGAAGCAGTCTTCTTCTGCTAGTTGAATTAATGATGATGTTTCCGTAATTAAGCAA 1578
QY 1004 ATAAATTT 1011
DB 1579 ATAAATTT 1586
```

RESULT 13  
US-10-847-918-14  
Sequence 14, Application US/10847918  
Publication No. US20050119210A1  
GENERAL INFORMATION:  
APPLICANT: Wyeth  
APPLICANT: Be, Xiaobing  
APPLICANT: Liu, Wei  
APPLICANT: Slonim, Donna  
APPLICANT: Howes, Steve  
TITLE OF INVENTION: Compositions and Methods for Diagnosing and Treating Cancers  
FILE REFERENCE: 031896-026000 (JMI01264)  
CURRENT APPLICATION NUMBER: US/10/847,918  
PRIOR FILING DATE: 2004-05-19  
PRIORITY FILING DATE: 2003-05-20  
NUMBER OF SEQ ID NOS: 14937  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 14  
LENGTH: 1586  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-847-918-14

Query Match 81.6%; Score 838; DB 9; Length 1586;  
Best Local Similarity 88.4%; Pred. No. 1.8e-237;  
Matches 962; Conservative 0; Mismatches 30; Indels 96; Gaps 1;

```
QY 20 AAATCAAGTATTTGTCCTCCGGAACAAAGCTTATCATACAGATTAACAACCTTGATGACG 79
DB 499 ATAAAGAACTCTTTGGCAGTGAACAACCTTGTTGTACAGATTAACAACCTTGATGACG 558
QY 80 ATGTTTCCCAAGCCCACTATTTTCTTCTCAATTGCTGAACAAAGCTCCAGAGG 139
DB 559 ATGTTTCCCAAGCCCACTATTTTCTTCTCAATTGCTGAACAAAGCTCCAGAGG 618
QY 140 CTGAACATACCTTTGCTTCTGAGAAATTTTCCCGATGTTATTAAGATGATGACG 199
DB 619 CTGAACATATCTTTGCTTCTGAGAAATTTTCCCGATGTTATTAAGATGATGACG 678
QY 200 AAGAAAGAGAGACAAACAATTCGGATCCAGAGGGAGAAACCATGAAGACTAAGC 259
DB 679 AAGAAAGAGAGACAAACAATTCGGATCCAGAGGGAGAAACCATGAAGACTAAGC 738
QY 260 ACACATACATGAATTTTACCTGCTTACCGTGCAGAAAGTCACTGACAAAGACACA 319
DB 739 ACACATACATGAATTTTACCTGCTTACCGTGCAGAAAGTCACTGACAAAGACACA 798
QY 320 GATGATCGTGAACATGAGAAATTAATAACGGAGTTGATCAAGAAATTTCTTCTC 379
DB 799 GATGATCGTGAACATGAGAAATTAATAACGGAGTTGATCAAGAAATTTCTTCTC 858
QY 380 CAATTAAGAC----- 389
DB 859 CAATTAAGACAGATGTACACAGATGATCCCAAGACAGTTATTCAAAAGATGCAATG 918
```



RESULT 15  
US-10-357-930-24142/c  
; Sequence 24142, Application US/10357930  
; Publication No. US20040259086A1  
; GENERAL INFORMATION:  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Endege, Wilson  
; APPLICANT: Monahan, John  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF  
; TITLE OF INVENTION: HUMAN PROSTATE CANCER  
; FILE REFERENCE: MRI-007BCN  
; CURRENT APPLICATION NUMBER: US/10/357,930  
; PRIOR FILING DATE: 2003-02-04  
; PRIOR APPLICATION NUMBER: 09/785,276  
; PRIOR FILING DATE: 2003-02-16  
; PRIOR APPLICATION NUMBER: 60/183,319  
; PRIOR FILING DATE: 2000-02-17  
; PRIOR APPLICATION NUMBER: 60/189,862  
; PRIOR FILING DATE: 2000-03-16  
; PRIOR APPLICATION NUMBER: 60/207,454  
; PRIOR FILING DATE: 2000-05-25  
; PRIOR APPLICATION NUMBER: 60/211,314  
; PRIOR FILING DATE: 2000-06-09  
; PRIOR APPLICATION NUMBER: 60/219,007  
; PRIOR FILING DATE: 2000-07-18  
; PRIOR APPLICATION NUMBER: 60/255,281  
; PRIOR FILING DATE: 2000-12-13  
; NUMBER OF SEQ ID NOS: 62232  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 24142  
; LENGTH: 1162  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-357-930-24142

Query Match 77.5%; Score 795.8; DB 8; Length 1162;  
Best Local Similarity 99.7%; Pred. No. 5.3e-225;  
Matches 797; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 229 TCCGAGGAGGGAACCATTAAGCATTAAGACATATCATGAATTTAGCTGGTTAACG 288  
DB 1162 TCCGAGGAGGGAACCATTAAGCATTAAGACATATCATGAATTTAGCTGGTTAACG 1103  
QY 289 GTGCAGAAAGTCACTGCAAGAAACAGATGTATGTCAGACATGAGATTAATAA 348  
DB 1102 GTGCAGAAAGTCACTGCAAGAAACAGATGTATGTCAGACATGAGATTAATAA 1043  
QY 349 AACGAGTTGATCAAGAAATTAATCTTCTCCAAATAAAGCGATGTCATCAATGGAT 408  
DB 1042 AACGAGTTGATCAAGAAATTAATCTTCTCCAAATAAAGCGATGTCATCAATGGAT 983  
QY 409 CCCAAAGCAATTTGTTCAAAAGATGCAATATACACTACTGCTGCACTCAACAAACCC 468  
DB 982 CCCAAAGCAATTTGTTCAAAAGATGCAATATACACTACTGCTGCACTCAACAAACCC 923  
QY 469 TCTGATATTACATTAATCTCTCTGCTCTCAAGAGTGTGTATTTTGCCATCATC 528  
DB 922 TCTGATATTACATTAATCTCTCTGCTCTCAAGAGTGTGTATTTTGCCATCATC 863  
QY 529 ACCTGCTGTCTGCTTAGAAGAAAGGCTTTCTGCTGCAATGAGAGAAATCATTAACAGC 588  
DB 862 ACCTGCTGTCTGCTTAGAAGAAAGGCTTTCTGCTGCAATGAGAGAAATCATTAACAGC 803  
QY 589 GTGGACAAGAGAGGCACTTTTCTCTCATCGTTATTTGCTCTAGAACGCTTTTGAGG 648  
DB 802 GTGGACAAGAGAGGCACTTTTCTCTCATCGTTATTTGCTCTAGAACGCTTTTGAGG 743  
QY 649 ATCTAGTGGGCTTCTTCTTCTGAGTTTGAGGCATTTTCAGTTCTCATGTGTACTATTCT 708  
DB 742 ATCTAGTGGGCTTCTTCTTCTGAGTTTGAGGCATTTTCAGTTCTCATGTGTACTATTCT 683

QY 709 ATCATTTTGTATTAAGGTTTTCAAACCAAGTGGGCAACAGAGAACTCTGTATATA 768  
DB 682 ATCATTTTGTATTAAGGTTTTCAAACCAAGTGGGCAACAGAGAACTCTGTATATA 623  
QY 769 ACAATGAGATAGCCAGCGGATCTCCAGACCAATCTCTCATGTTTTCACAGCTCC 828  
DB 622 ACAATGAGATAGCCAGCGGATCTCCAGACCAATCTCTCATGTTTTCACAGCTCC 563  
QY 829 TCCAGCCAAACCAATAGCGGCTGCTATAGTATAGACATCTGCGGCTTCTAGCCTTGTG 888  
DB 562 TCCAGCCAAACCAATAGCGGCTGCTATAGTATAGACATCTGCGGCTTCTAGCCTTGTG 503  
QY 889 CCTCTTATAGTCTTCTTATCAATTAACCTGCTGAAAGCTTTCATTTTACAGCCCTG 948  
DB 502 CCTCTTATAGTCTTCTTATCAATTAACCTGCTGAAAGCTTTCATTTTACAGCCCTG 443  
QY 949 AAGCAGCTTCTTGTAGTTGAATTAATGATGATGTTTTCGTAATTAAGCAAAATATA 1008  
DB 442 AAGCAGCTTCTTGTAGTTGAATTAATGATGATGTTTTCGTAATTAAGCAAAATATA 383  
QY 1009 TTTAAAAAATGAAGATT 1027  
DB 382 TTTAAAAAATGAAGATT 364

Search completed: December 11, 2005, 01:30:51  
Job time : 895 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2005, 20:30:37 ; Search time 230 Seconds

(without alignments)  
7937.199 Million cell updates/sec

Title: US-10-031-158b-13

Perfect score: 1027

Sequence: 1 GGGCAAGAGCTGGGCAAGAAAAA.....atttaaaatgaaagct 1027

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

Issued Patents NA: \*  
1: /cgn2\_6/prodata1/ina/1 COMB.seq: \*  
2: /cgn2\_6/prodata1/ina/5 COMB.seq: \*  
3: /cgn2\_6/prodata1/ina/6A COMB.seq: \*  
4: /cgn2\_6/prodata1/ina/6B COMB.seq: \*  
5: /cgn2\_6/prodata1/ina/H COMB.seq: \*  
6: /cgn2\_6/prodata1/ina/PCTUS COMB.seq: \*  
7: /cgn2\_6/prodata1/ina/PP COMB.seq: \*  
8: /cgn2\_6/prodata1/ina/RE COMB.seq: \*  
9: /cgn2\_6/prodata1/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	516.4	50.3	1080	9 5260223-3	Patent No. 5260223
2	449.2	43.7	825	2 US-08-256-964A-18	Sequence 18, Appl
3	444	43.2	511	3 US-09-020-956-72	Sequence 72, Appl
4	444	43.2	511	3 US-09-030-607-72	Sequence 72, Appl
5	444	43.2	511	3 US-09-439-313-72	Sequence 72, Appl
6	444	43.2	511	3 US-09-352-616A-72	Sequence 72, Appl
7	444	43.2	511	3 US-09-232-149A-72	Sequence 72, Appl
8	444	43.2	511	3 US-09-159-812-72	Sequence 72, Appl
9	444	43.2	511	3 US-09-636-215-72	Sequence 72, Appl
10	444	43.2	511	3 US-09-166A-72	Sequence 72, Appl
11	444	43.2	511	3 US-09-115-453-72	Sequence 72, Appl
12	444	43.2	511	3 US-09-688-489-72	Sequence 72, Appl
13	444	43.2	511	3 US-09-679-426-72	Sequence 72, Appl
14	444	43.2	511	3 US-09-759-143-72	Sequence 72, Appl
15	444	43.2	511	3 US-09-651-236-72	Sequence 72, Appl
16	444	43.2	511	3 US-09-030-606-72	Sequence 72, Appl
17	444	43.2	511	3 US-09-657-279-72	Sequence 72, Appl
18	444	43.2	511	3 US-10-012-896-72	Sequence 72, Appl
19	407	39.6	467	3 US-09-020-956-75	Sequence 75, Appl
20	407	39.6	467	3 US-09-030-607-75	Sequence 75, Appl
21	407	39.6	467	3 US-09-439-313-75	Sequence 75, Appl
22	407	39.6	467	3 US-09-352-616A-75	Sequence 75, Appl
23	407	39.6	467	3 US-09-232-149A-75	Sequence 75, Appl
24	407	39.6	467	3 US-09-159-812-75	Sequence 75, Appl

25	407	39.6	467	3 US-09-636-215-75	Sequence 75, Appl
26	407	39.6	467	3 US-09-685-166A-75	Sequence 75, Appl
27	407	39.6	467	3 US-09-115-453-75	Sequence 75, Appl
28	407	39.6	467	3 US-09-688-489-75	Sequence 75, Appl
29	407	39.6	467	3 US-09-679-426-75	Sequence 75, Appl
30	407	39.6	467	3 US-09-759-143-75	Sequence 75, Appl
31	407	39.6	467	3 US-09-651-236-75	Sequence 75, Appl
32	407	39.6	467	3 US-09-030-606-75	Sequence 75, Appl
33	407	39.6	467	3 US-09-657-279-75	Sequence 75, Appl
34	407	39.6	467	3 US-10-012-896-75	Sequence 75, Appl
35	404.4	39.4	426	3 US-09-513-999C-679	Sequence 679, Appl
36	250.6	24.4	281	3 US-09-439-313-433	Sequence 433, Appl
37	250.6	24.4	281	3 US-09-352-616A-433	Sequence 433, Appl
38	250.6	24.4	281	3 US-09-636-215-433	Sequence 433, Appl
39	250.6	24.4	281	3 US-09-685-166A-433	Sequence 433, Appl
40	250.6	24.4	281	3 US-09-679-426-433	Sequence 433, Appl
41	250.6	24.4	281	3 US-09-759-143-433	Sequence 433, Appl
42	250.6	24.4	281	3 US-09-651-236-433	Sequence 433, Appl
43	250.6	24.4	281	3 US-09-657-279-433	Sequence 433, Appl
44	250.6	24.4	281	3 US-10-012-896-433	Sequence 433, Appl
45	247.6	24.1	460	3 US-09-513-999C-8197	Sequence 8197, Appl

#### ALIGNMENTS

RESULT 1					
5260223-3					
Patent No. 5260223					
; APPLICANT: BRENNER, MICHAEL B.; STROMINGER, JACK L.; SEIDMAN, JOHN G.; IP, STEPHEN H.; KRANGE, MICHAEL S.					
; TITLE OF INVENTION: METHODS FOR DETECTION OF HUMAN GAMMA					
; T CELL RECEPTOR					
; NUMBER OF SEQUENCES: 4					
; CURRENT APPLICATION DATA:					
; APPLICATION NUMBER: US/07/187,698					
; FILING DATE: 29-APR-1988					
; PRIOR APPLICATION DATA:					
; APPLICATION NUMBER: 115,256					
; FILING DATE: 29-OCT-1987					
; APPLICATION NUMBER: 16,252					
; FILING DATE: 19-FEB-1987					
; APPLICATION NUMBER: 882,100					
; FILING DATE: 03-JUL-1986					
; SEQ ID NO:3:					
; LENGTH: 1080					
5260223-3					
Query Match					
Best Local Similarity 50.3%; Score 516.4; DB 9; Length 1080;					
Best Local Similarity 88.9%; Pred. No. 8.4e-150;					
Matches 590; Conservative 0; Mismatches 26; Indels 48; Gaps 1;					
QY	25	AAAGTATTTGGTCCCGAACAAGCTTATCATTAAGATTAACAACTTGATGAGATGTT	84		
DB	403	AAACTTTTGGAGTGGAACAACACTTGTGCAGATTAACAACCTGATGAGATGTT	462		
QY	85	TCCCCCAAGCCCACTATTTTCTTCTTCAATGCTGGAACAAAGCTCCGAAGGCTGGA	144		
DB	463	TCCCCCAAGCCCACTATTTTCTTCTTCAATGCTGGAACAAAGCTCCGAAGGCTGGA	522		
QY	145	ACATACCTTGTGTTCTTGGAATTTTCCCTGATGTTATTAAGATTAATGCTGCAAGAA	204		
DB	523	ACATACCTTGTGTTCTTGGAATTTTCCCTGATGTTATTAAGATTAATGCTGCAAGAA	582		
QY	205	AAAGAGCAACGATCTTGAGATCCAGAGAGGGAACACATGAAGACTTAAGACACA	264		
DB	583	AAAGAGCAACGATCTTGAGATCCAGAGAGGGAACACATGAAGACTTAAGACACA	642		
QY	265	TACATGAATTTAGCTGCTTAACGCTGCAAGAAATTCCTGACAAAGAACACATGT	324		
DB	643	TACATGAATTTAGCTGCTTAACGCTGCAAGAAATTCCTGACAAAGAACACATGT	702		
QY	325	ATGTCAGACATGAGATTAATAAAACGAGGTGATCAAGAAATTAATCTTCTCCATA	384		

Db	703	ATCGTCAGCATGTAGAAATPAATPAATAAAGCGAATTGATCAAGAAATTATCTTTGCTCCATA	762
QY	385	AAGC-----GGATGTC	396
Db	763	AAGCAGATGTCAACACAGTGATCCCAATATCAATTATTCAAAGATGCAAATGATGTC	822
QY	397	ATCACAATGATATCCCAAGACATTTGTTCAAAAGATGCAATGATACCTACTGTGAG	456
Db	823	ATCACAATGATATCCCAAGACATTTGTTCAAAAGATGCAATGATACCTACTGTGAG	882
QY	457	CTCACAAACACTGTGCATATTACATGTACTCTCTCTCTCTCAAGATGTGTCTAT	516
Db	883	CTCAGAAACACTGTGCATATTACACGTACTCTCTCTCTCTCAAGATGTGTCTAT	942
QY	517	TTTGCCATCATCACTGCTGTCTGTTAGAAAGACGGCTTTCTGTGCAATGAGAGAAA	576
Db	943	TTTGCCATCATCACTGCTGTCTGTTAGAAAGACGGCTTTCTGTGCAATGAGAGAAA	1002
QY	577	TCATPACAGACGGTGGCAAGAGAGGCATCTTTTCCATCGGTTATGTGCCCTAGAG	636
Db	1003	TCATPACAGACGGTGGCAAGAGAGGCATCTTTTCCATCGGTTATGTGCCCTAGAG	1062
QY	637	CGTC 640	
Db	1063	CGTC 1066	

```

? NAME/KEY: CDS
? LOCATION: 1..825
? PUBLICATION INFORMATION:
? DOCUMENT NUMBER: WO 94/12648
? FILING DATE: 25-NOV-1993
? PUBLICATION DATE: 09-JUN-1994
US-08-256-964A-18

Query Match 43.7%; Score 449.2; DB 2; Length 825;
Best Local Similarity 99.3%; Pred. No. 5.5e-129;
Matches 451; Conservative 0; Mismatches 3; Indels 0; Gaps 0

QY 3 GCAAGAGTGGGCAAAAAATCAAGATGTTTGGTCCGGAAACAAGCTTATCATTCACGA 62
Db 369 GCAAGAGTGGGCAAAAAATCAAGATGTTTGGTCCGGAAACAAGCTTATCATTCACGA 428

QY 63 TAAACAACTTGATGCAGATGTTTCCCCCAAGCCCACTATTTTCTTCTTCAATTGCTGA 122
Db 429 TAAACAACTTGATGCAGATGTTTCCCCCAAGCCCACTATTTTCTTCTTCAATTGCTGA 488

QY 123 AACCAAGCTCCAGAAAGCTGGAACAATACCTTTGCTTCTTGAGAAATTTTCCCTGATGT 182
Db 489 AACCAAGCTCCAGAAAGCTGGAACAATACCTTTGCTTCTTGAGAAATTTTCCCTGATGT 548

QY 183 TATTAGATCATTTGGCAAGAAAAGAGCAACGATTCCTGGGATCCCAAGAGGGGA 242

```

```

1      RESULT 2
2      US-08-256-964A-18
3      : Sequence 18, Application US/08256964A
4      : Patent No. 5723309
5      : GENERAL INFORMATION:
6      : APPLICANT: BONNEVILLE, Marc
7      : TITLE OF INVENTION: PRODUCTION OF SUBUNITS OF SOLUBLE T
8      : TITLE OF INVENTION: RECEPTORS BY CO-TRANSECTION, USES OF THE PRODUCTS THUS
9      : TITLE OF INVENTION: OBTAINED
10     : NUMBER OF SEQUENCES: 19
11     : CORRESPONDENCE ADDRESS:
12     : ADDRESSEE: Young & Thompson
13     : STREET: 745 South 23rd Street
14     : CITY: Arlington
15     : STATE: VA
16     : COUNTRY: USA
17     : ZIP: 22202
18     : COMPUTER READABLE FORM:
19     : MEDIUM TYPE: Floppy disk
20     : COMPUTER: IBM PC compatible
21     : OPERATING SYSTEM: PC-DOS/MS-DOS
22     : SOFTWARE: PatentIn Release #1.0, Version #1.25
23     : CURRENT APPLICATION DATA:
24     : APPLICATION NUMBER: US/08/256,964A
25     : FILING DATE: 14-SEP-1994
26     : CLASSIFICATION: 435
27     : PRIOR APPLICATION DATA:
28     : APPLICATION NUMBER: FR 92 14203
29     : FILING DATE: 25-NOV-1992
30     : ATTORNEY/AGENT INFORMATION:
31     : NAME: PATCH, Andrew J.
32     : REGISTRATION NUMBER: Reg. No. 5723309 32,925
33     : REFERENCE/DOCKET NUMBER: BR 94/449
34     : TELECOMMUNICATION INFORMATION:
35     : TELEPHONE: 703/521-2297
36     : TELEFAX: 703/685-0573
37     : TELEX: 248425
38     : INFORMATION FOR SEQ. ID NO: 18:
39     : SEQUENCE CHARACTERISTICS:
40     : LENGTH: 825 base pairs
41     : TYPE: nucleic acid
42     : STRANDEDNESS: single
43     : TOPOLOGY: linear
44     : MOLECULE TYPE: DNA (genomic)
45     : FEATURE:

```

```

1  RESULT 3
2  US-09-020-956-72/C
3  ; Sequence 72, Application US/09020956
4  ; Patent No. 6261562
5  ; GENERAL INFORMATION:
6  ; APPLICANT: Xu, Jiangchun
7  ; APPLICANT: Dillin, Davin C.
8  ; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FOR
9  ; NUMBER OF SEQUENCES: 178
10 ; CORRESPONDENCE ADDRESS:
11 ; ADDRESSEE: SEED and BERRY LLP
12 ; STREET: 6300 Columbia Center, 701 Fifth Avenue
13 ; CITY: Seattle
14 ; STATE: WA
15 ; COUNTRY: USA
16 ; ZIP: 98104
17 ;
18 ; COMPUTER READABLE FORM:
19 ; MEDIUM TYPE: Floppy disk
20 ; COMPUTER: IBM PC compatible
21 ; OPERATING SYSTEM: PC-DOS/MS-DOS
22 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
23 ; CURRENT APPLICATION DATA:
24 ; APPLICATION NUMBER: US/09/020,956
25 ; FILING DATE: 09-FEB-1998
26 ; CLASSIFICATION:
27 ; ATTORNEY/AGENT INFORMATION:
28 ; NAME: MAKI, David J.
29 ; REGISTRATION NUMBER: 31,392

```

REFERENCE/DOCKET NUMBER: 210121.427C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031 72:  
INFORMATION FOR SEQ ID NO: 72:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 511 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
US-09-020-956-72

Query Match 43.2%; Score 444; DB 3; Length 511;  
Best Local Similarity 91.3%; Pred. No. 1.7e-127;  
Matches 462; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 492 CCGCTCTCAAGAGTGTGCTATTTTGGCATCATACCTGCTGCTTGAAGAAC 551  
DB 506 CNGCTTCTCAAGAGGCGGTGTATTTTNGCATCATACCTGCTTGAAGAAC 447  
QY 552 GCGTTTCTGCTGAATGAGAGAAATCATTAACAGCGGTGCAAGAGAGCCATCTTT 611  
DB 446 GGTTTTNGCTGCAATGAGAGAAATCATTAACAGCGGTGCAAGAGAGCCATCTTT 387  
QY 612 CCGTATCGTTATTTGCTTGAAGAGCTCTTCTGAGGATTTAGTTGGCTTTCTTCTGG 671  
DB 386 CTTTCATCGTTATTTGCTTGAAGAGCTTCTGAGGATTTAGTTGGCTTTCTTCTGG 327  
QY 672 GTTGGGCAATTCAGTCTCATGTGTGTAATCTATCATATTGTATTAAGGTTTC 731  
DB 326 GTTGGGCAATTCAGTCTCATGTGTGTAATCTATCATATTGTATTAAGGTTTC 267  
QY 732 AAACAGTGGGACACAGAGAACTCACTGTATTAACATGAGAAATGAGAGCGGCGA 791  
DB 266 AAACAGTGGGACACAGAGAACTCACTGTATTAACATGAGAAATGAGAGCGGCGA 207  
QY 792 TCTCCAGCAACATCTCTCATATGTTTCCACAGCTCTCCAGCCAAATAGCGCCT 851  
DB 206 TCTCCAGCAACATCTCTCATATGTTTCCACAGCTCTCCAGCCAAATAGCGCCT 147  
QY 852 GCTATGTGAGACATCTGCGGCTTCTAGCTTCTCTTATGTTCTTTATCA 911  
DB 146 GATTAATGTAGACATCTGCGGCTTCTAGCTTCTCTTATGTTCTTTATCA 87  
QY 912 GATTAATGTGAGACATCTTCAATTTTACAGCCCTGAGAGAGCTTCTTGTAGTTGA 971  
DB 86 GATTAATGTGAGACATCTTCAATTTTACAGCCCTGAGAGAGCTTCTTGTAGTTGA 27  
QY 972 ATTATGTGTGTGTTTTCGGTATA 997  
DB 26 ATTATGTGTGTGTTTTCGGTATA 1

RESULT 4  
US-09-030-607-72/c  
; Sequence 72, Application US/09030607  
; Patent No. 6262245  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiaqun  
; APPLICANT: Dillon, David C.  
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO  
; NUMBER OF SEQUENCES: 224  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98104  
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/030,607  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.427C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031 72:  
INFORMATION FOR SEQ ID NO: 72:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 511 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
US-09-030-607-72

Query Match 43.2%; Score 444; DB 3; Length 511;  
Best Local Similarity 91.3%; Pred. No. 1.7e-127;  
Matches 462; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 492 CCGCTCTCAAGAGTGTGCTATTTTGGCATCATACCTGCTGCTTGAAGAAC 551  
DB 506 CNGCTTCTCAAGAGGCGGTGTATTTTNGCATCATACCTGCTTGAAGAAC 447  
QY 552 GCGTTTCTGCTGAATGAGAGAAATCATTAACAGCGGTGCAAGAGAGCCATCTTT 611  
DB 446 GGTTTTNGCTGCAATGAGAGAAATCATTAACAGCGGTGCAAGAGAGCCATCTTT 387  
QY 612 CCGTATCGTTATTTGCTTGAAGAGCTCTTCTGAGGATTTAGTTGGCTTTCTTCTGG 671  
DB 386 CTTTCATCGTTATTTGCTTGAAGAGCTTCTGAGGATTTAGTTGGCTTTCTTCTGG 327  
QY 672 GTTGGGCAATTCAGTCTCATGTGTGTAATCTATCATATTGTATTAAGGTTTC 731  
DB 326 GTTGGGCAATTCAGTCTCATGTGTGTAATCTATCATATTGTATTAAGGTTTC 267  
QY 732 AAACAGTGGGACACAGAGAACTCACTGTATTAACATGAGAAATGAGAGCGGCGA 791  
DB 266 AAACAGTGGGACACAGAGAACTCACTGTATTAACATGAGAAATGAGAGCGGCGA 207  
QY 792 TCTCCAGCAACATCTCTCATATGTTTCCACAGCTCTCCAGCCAAATAGCGCCT 851  
DB 206 TCTCCAGCAACATCTCTCATATGTTTCCACAGCTCTCCAGCCAAATAGCGCCT 147  
QY 852 GCTATGTGAGACATCTGCGGCTTCTAGCTTCTCTTATGTTCTTTATCA 911  
DB 146 GATTAATGTAGACATCTGCGGCTTCTAGCTTCTCTTATGTTCTTTATCA 87  
QY 912 GATTAATGTGAGACATCTTCAATTTTACAGCCCTGAGAGAGCTTCTTGTAGTTGA 971  
DB 86 GATTAATGTGAGACATCTTCAATTTTACAGCCCTGAGAGAGCTTCTTGTAGTTGA 27  
QY 972 ATTATGTGTGTGTTTTCGGTATA 997  
DB 26 ATTATGTGTGTGTTTTCGGTATA 1

RESULT 5  
US-09-439-313-72/c  
; Sequence 72, Application US/09439313  
; Patent No. 6329505  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiaqun

APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan Louise  
APPLICANT: Jiang, Yugu  
APPLICANT: Reed, Steven G.  
APPLICANT: Kaios, Michael  
APPLICANT: Fanger, Gary  
APPLICANT: Retter, Mark  
APPLICANT: Soik, John  
APPLICANT: Day, Craig  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND  
FILE REFERENCE: 210121.427C9  
CURRENT APPLICATION NUMBER: US/09/439,313  
CURRENT FILING DATE: 1999-11-12  
NUMBER OF SEQ ID NOS: 575  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 72  
LENGTH: 511  
TYPE: DNA  
ORGANISM: Homo sapien  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)\_(511)  
OTHER INFORMATION: n = A,T,C or G  
US-09-439-313-72

Query Match 43.2%; Score 444; DB 3; Length 511;  
Best Local Similarity 91.3%; Pred. No. 1.7e-127;  
Matches 462; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 492 CCGCTCTCTCAAGAGTGTGTCTATTTTGGCCATCATCCTGCTGTCTTGAAGAAC 551  
Db 506 CNGGCTCTCAAGAGGAGTGTATTTTGGCCATCATCCTGCTGTCTTGAAGAACG 447  
Qy 552 GCGTTTGTGTCATGAGAGAAATCATACAGCGGTGACAGAGGCGCATCTTTT 611  
Db 446 GGTITNTGTCGATGAGAGAAATCATACAGCGGAGAGAGGCGCATCTTTT 387  
Qy 612 CCTCATCGTTATTTGCTCCCTAGAAAGCTCTTCTGAGATCTAGTTGGCTTTCTTTCTGG 671  
Db 386 CTTATCGTTATTTGCTCCCTAGAAAGCTTTCTGAGATCTAGTTGGCTTTCTTTCTGG 327  
Qy 672 GTTTGGGCAATTTCAAGTTCTCATGTGTGTACTATTTATCATATTGTATTAACGTTTTT 731  
Db 326 GTTTGGGCAATTTCAAGTTCTCATGTGTGTACTATTTATCATATTGTATTAACGTTTTT 267  
Qy 732 AAACCAAGTGGGACACAGAGAACTCTCTGTATTAACATGAGAAATAGCCAGCGGA 791  
Db 266 AAACCAAGTGGGACACAGAGAACTCTCTGTATTAACATGAGAAATAGCCAGCGGA 207  
Qy 792 TCTCCAGACCAATCTCTCATGTTTTTCCAGAGCTCTCCAGCAACCAATAGCGCT 851  
Db 206 TTTCCAGACCAATCTCTCATGTTTTTCCAGAGCTCTCCAGCAACCAATAGCGCT 147  
Qy 852 GCTATAGTGAACATCTCTGCGGCTTCTAGCGCTGTCTCTTATGTTCTTAAATCA 911  
Db 146 GNTATAGTGAACATCTCTGCGGCTTCTAGCGCTGTCTCTTATGTTCTTAAATCA 87  
Qy 912 GATACTGCTGGAAGCTTTTCAATTTTACAGCCCTGAGAGAGCTTTCTTGTCTAGTTGA 971  
Db 86 GATACTGCTGGAAGCTTTTCAATTTTACAGCCCTGAGAGAGTNTTCTTGTAGTTGA 27  
Qy 972 ATTATGTGTGTGTTTTCCGTAATA 997  
Db 26 ATTATGTGTGTGTTTTCCGTAATA 1

RESULT 6  
US-09-352-616A-72/c  
Sequence 72, Application US/09352616A  
Patent No. 6395278  
GENERAL INFORMATION:

APPLICANT: Dillon, Davin C.  
APPLICANT: Harlocker, Susan Louise  
APPLICANT: Jiang, Yugu  
APPLICANT: Xu, Jiangchun  
APPLICANT: Mitcham, Jennifer Lynn  
TITLE OF INVENTION: OF PROSTATE CANCER AND METHODS FOR THEIR USE  
FILE REFERENCE: 210121.427C8  
CURRENT APPLICATION NUMBER: US/09/352,616A  
CURRENT FILING DATE: 1999-07-13  
NUMBER OF SEQ ID NOS: 472  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 72  
LENGTH: 511  
TYPE: DNA  
ORGANISM: Homo sapien  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)...(511)  
OTHER INFORMATION: n = A,T,C or G  
US-09-352-616A-72

Query Match 43.2%; Score 444; DB 3; Length 511;  
Best Local Similarity 91.3%; Pred. No. 1.7e-127;  
Matches 462; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 492 CCGCTCTCTCAAGAGTGTGTCTATTTTGGCCATCATCCTGCTGTCTTGAAGAAC 551  
Db 506 CNGGCTCTCAAGAGGAGTGTATTTTGGCCATCATCCTGCTGTCTTGAAGAACG 447  
Qy 552 GCGTTTGTGTCATGAGAGAAATCATACAGCGGTGACAGAGGCGCATCTTTT 611  
Db 446 GGTITNTGTCGATGAGAGAAATCATACAGCGGAGAGAGGCGCATCTTTT 387  
Qy 612 CCTCATCGTTATTTGCTCCCTAGAAAGCTCTTCTGAGATCTAGTTGGCTTTCTTTCTGG 671  
Db 386 CTTATCGTTATTTGCTCCCTAGAAAGCTTTCTGAGATCTAGTTGGCTTTCTTTCTGG 327  
Qy 672 GTTTGGGCAATTTCAAGTTCTCATGTGTGTACTATTTATCATATTGTATTAACGTTTTT 731  
Db 326 GTTTGGGCAATTTCAAGTTCTCATGTGTGTACTATTTATCATATTGTATTAACGTTTTT 267  
Qy 732 AAACCAAGTGGGACACAGAGAACTCTCTGTATTAACATGAGAAATAGCCAGCGGA 791  
Db 266 AAACCAAGTGGGACACAGAGAACTCTCTGTATTAACATGAGAAATAGCCAGCGGA 207  
Qy 792 TCTCCAGACCAATCTCTCATGTTTTTCCAGAGCTCTCCAGCAACCAATAGCGCT 851  
Db 206 TTTCCAGACCAATCTCTCATGTTTTTCCAGAGCTCTCCAGCAACCAATAGCGCT 147  
Qy 852 GCTATAGTGAACATCTCTGCGGCTTCTAGCGCTGTCTCTTATGTTCTTAAATCA 911  
Db 146 GNTATAGTGAACATCTCTGCGGCTTCTAGCGCTGTCTCTTATGTTCTTAAATCA 87  
Qy 912 GATACTGCTGGAAGCTTTTCAATTTTACAGCCCTGAGAGAGCTTTCTTGTCTAGTTGA 971  
Db 86 GATACTGCTGGAAGCTTTTCAATTTTACAGCCCTGAGAGAGTNTTCTTGTAGTTGA 27  
Qy 972 ATTATGTGTGTGTTTTCCGTAATA 997  
Db 26 ATTATGTGTGTGTTTTCCGTAATA 1

RESULT 7  
US-09-232-149A-72/c  
Sequence 72, Application US/09232149A  
Patent No. 6465611  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer Lynn  
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE  
CANCER AND METHODS FOR THEIR USE

FILE REFERENCE: 210121.427C6  
CURRENT APPLICATION NUMBER: US/09/232.149A  
CURRENT FILING DATE: 1999-01-15  
NUMBER OF SEQ ID NOS: 338  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO 72  
LENGTH: 511  
TYPE: DNA  
ORGANISM: Homo sapien  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (1)...(511)  
OTHER INFORMATION: n = A,T,C or G  
US-09-232-149A-72

Query Match 43.2%; Score 444; DB 3; Length 511;  
Best Local Similarity 91.3%; Pred. No. 1,7e-127;  
Matches 462; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

```
QY 492 CCTGCTCTCAAGAGTGTGTCTATTGTCATCATACCTGCTGTCTGTTAGAAAGAC 551
DB 506 CNGGTCCTCAAGAGGGGGGTATTTTNGCCATCATCCTGNGTTGCTTAGAAGACG 447
QY 552 GCGTTTCTGCGAATGAGAGAAATCATACAGACGGTGGCAAGAGAGGCCATCTTTT 611
DB 446 GGTTTTNGCTGCAATGAGAGAAATCATACAGACGGGGGCAAGAGAGGCCATCTTTT 387
QY 612 CCTCATCGGTATATGTCCTTAGAAGCGCTCTTGAGAGATCTAGTGGGCTTTCTTCTGG 671
DB 386 CTTCATCGGTATATGTCCTTAGAAGCGCTCTTGAGAGATCTAGTGGGCTTTCTTCTGG 327
QY 672 GTTTGGGCCATTTCAGTTCATGTGTGTAATCTATCTATCTATTAATTAAGGTTTTC 731
DB 326 GTTTGGGCCATTTCAGTTCATGTGTGTAATCTATCTATCTATTAATTAAGGTTTTC 267
QY 732 AAACGATGGGACACAGAGAACTCACTGTATTAACATGAGAAATGACGACGGCA 791
DB 266 AAACGATGGGACACAGAGAACTCACTGTATTAACATGAGAAATGACGACGGCA 207
QY 792 TCTCCAGACCAATCTCTCATATGTTTCCACAGCTCTCTCCAGCCAAACCAATAGCGCT 851
DB 206 TCTCCAGACCAATCTCTCATATGTTTCCACAGCTCTCTCCAGCCAAACCAATAGCGCT 147
QY 852 GCTATAGTGAACATCTCTGCGGCTTCTAGCCTTGTCTCTGATGTTCTTTATCA 911
DB 146 GATATAGTGAACATCTCTGCGGCTTCTAGCCTTGTCTCTGATGTTCTTTATCA 87
QY 912 GATACCTGCTGGAAGCCTTTCATTTTACAGCCCTGAGAGAGTCTTCTGCTAGTTGA 971
DB 86 GATACCTGCTGGAAGCCTTTCATTTTACAGCCCTGAGAGAGTCTTCTGATGTTGA 27
QY 972 ATTATGTGTGTGTGTTTTCGTAATA 997
DB 26 ATTATGTGTGTGTGTTTTCGTAATA 1
```

RESULT 8  
US-09-159-812-72/c  
Sequence 72, Application US/09159812A  
Patent No. 6613872  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, Davin C.  
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF  
FILE REFERENCE: 210121.428C5  
CURRENT APPLICATION NUMBER: US/09/159, 812A  
CURRENT FILING DATE: 1998-09-23  
NUMBER OF SEQ ID NOS: 306  
SOFTWARE: FASTSEQ for Windows Version 3.0  
SEQ ID NO 72  
LENGTH: 511  
TYPE: DNA

ORGANISM: Homo sapien  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (1)...(511)  
OTHER INFORMATION: n = A,T,C or G  
US-09-159-812-72

Query Match 43.2%; Score 444; DB 3; Length 511;  
Best Local Similarity 91.3%; Pred. No. 1,7e-127;  
Matches 462; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

```
QY 492 CCTGCTCTCAAGAGTGTGTCTATTGTCATCATACCTGCTGTCTGTTAGAAAGAC 551
DB 506 CNGGTCCTCAAGAGGGGGGTATTTTNGCCATCATCCTGNGTTGCTTAGAAGACG 447
QY 552 GCGTTTCTGCGAATGAGAGAAATCATACAGACGGTGGCAAGAGAGGCCATCTTTT 611
DB 446 GGTTTTNGCTGCAATGAGAGAAATCATACAGACGGGGGCAAGAGAGGCCATCTTTT 387
QY 612 CCTCATCGGTATATGTCCTTAGAAGCGCTCTTGAGAGATCTAGTGGGCTTTCTTCTGG 671
DB 386 CTTCATCGGTATATGTCCTTAGAAGCGCTCTTGAGAGATCTAGTGGGCTTTCTTCTGG 327
QY 672 GTTTGGGCCATTTCAGTTCATGTGTGTAATCTATCTATCTATTAATTAAGGTTTTC 731
DB 326 GTTTGGGCCATTTCAGTTCATGTGTGTAATCTATCTATCTATTAATTAAGGTTTTC 267
QY 732 AAACGATGGGACACAGAGAACTCACTGTATTAACATGAGAAATGACGACGGCA 791
DB 266 AAACGATGGGACACAGAGAACTCACTGTATTAACATGAGAAATGACGACGGCA 207
QY 792 TCTCCAGACCAATCTCTCATATGTTTCCACAGCTCTCTCCAGCCAAACCAATAGCGCT 851
DB 206 TCTCCAGACCAATCTCTCATATGTTTCCACAGCTCTCTCCAGCCAAACCAATAGCGCT 147
QY 852 GCTATAGTGAACATCTCTGCGGCTTCTAGCCTTGTCTCTGATGTTCTTTATCA 911
DB 146 GATATAGTGAACATCTCTGCGGCTTCTAGCCTTGTCTCTGATGTTCTTTATCA 87
QY 912 GATACCTGCTGGAAGCCTTTCATTTTACAGCCCTGAGAGAGTCTTCTGCTAGTTGA 971
DB 86 GATACCTGCTGGAAGCCTTTCATTTTACAGCCCTGAGAGAGTCTTCTGATGTTGA 27
QY 972 ATTATGTGTGTGTGTTTTCGTAATA 997
DB 26 ATTATGTGTGTGTGTTTTCGTAATA 1
```

RESULT 9  
US-09-636-215-72/c  
Sequence 72, Application US/09636215  
Patent No. 6620922  
GENERAL INFORMATION:  
APPLICANT: Xu, Jiangchun  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Jiang, Yugu  
APPLICANT: Henderson, Robert A.  
APPLICANT: Kalos, Michael D.  
APPLICANT: Fanger, Gary R.  
APPLICANT: Retter, Marc W.  
APPLICANT: Stolk, John A.  
APPLICANT: Day, Craig H.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Darrick  
APPLICANT: Li, Samuel  
APPLICANT: Wang, Aijun  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Hepler, William  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.42717C17

;; CURRENT APPLICATION NUMBER: US/09/636,215  
;; CURRENT FILING DATE: 2000-08-10  
;; NUMBER OF SEQ ID NOS: 852  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 72  
;; LENGTH: 511  
;; TYPE: DNA  
;; ORGANISM: Homo sapien  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1)..(511)  
;; OTHER INFORMATION: n = A,T,C or G  
US-09-636-215-72

Query Match 43.2%; Score 444; DB 3; Length 511;  
Best Local Similarity 91.3%; Pred. No. 1,7e-127;  
Matches 462; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 492 CTTGCTCTCAAGAGTGTGTCATATTTTCCATCATCAGCTGTCTGCTTAGAAGAAC 551  
DB 506 CNGGCTCTCAAGAGGGGGGTATTTTNGCCATCATCAGCTGNGTTGCTTAGAAGAACG 447  
QY 552 GGCCTTCTGCTGCAATGAGAGAAATCATATACAGAGGTCAGAGAGGCAATCTTTT 611  
DB 446 GGTTCATGCTGCAATGAGAGAAATCATATACAGAGGTCAGAGAGGCAATCTTTT 387  
QY 612 CCTCATGCTATTTGTCCTTGAAGAGCTCTTGAAGATCTAGTTGGGCTTTCTTCTGG 671  
DB 386 CTTTCATGCTATTTGTCCTTGAAGAGCTTCTGAAGATCTAGTTGGGCTTTCTTCTGG 327  
QY 672 GTTTGGGCAATTCAGTTCTCATGTGTGACTATTTATCATATTTATTAACGTTTTC 731  
DB 326 GTTTGGGCAATTCAGTTCTCATGTGTGACTATTTATCATATTTATTAACGTTTTC 267  
QY 732 AAACCAAGGGGACACAGAGAACTCTGTAATTAACAATGAGAAATAGCCAGGCGA 791  
DB 266 AAACCAAGGGGACACAGAGAACTCTGTAATTAACAATGAGAAATAGCCAGGCGA 207  
QY 792 TCTCCAGACCAATCTCTCATGTTTTCACAGAGCTCTCCAGCCAAATAGCGCCT 851  
DB 206 TCTCCAGACCAATCTCTCATGTTTTCACAGAGCTCTCCAGCCAAATAGCGCCT 147  
QY 852 GCATAGTGTAGACATCTCGGCTTCTAGGCTTGTCTCTTAAATGCTTTTAAATCA 911  
DB 146 GNTATAGTGTAGACATCTCGGCTTCTAGGCTTGTCTCTTAAATGCTTTTAAATCA 87  
QY 912 GATAACGCTGGAAGGCTTTCATTTTACAGGCTGGAAGAGCTCTTGTGCTAGTTGA 971  
DB 86 GATAACGCTGGAAGGCTTTCATTTTACAGGCTGGAAGAGCTTGTGCTAGTTGA 27  
QY 972 ATTATGAGTGTGTTTCCGTAAATA 997  
DB 26 ATTATGAGTGTGTTTCCGTAAATA 1

RESULT 10  
US-09-685-166A-72/c  
; Sequence 72, Application US/09685166A  
; Patent No. 6630305  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, David C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Jiang, Yuqi  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Reiter, Marc W.  
; APPLICANT: Stolk, John A.  
; APPLICANT: Day, Craig H.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Derrick

;; APPLICANT: Li, Samuel  
;; APPLICANT: Wang, Aijun  
;; APPLICANT: Skeiky, Yasir A.W.  
;; APPLICANT: Hepler, William  
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
;; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
;; FILE REFERENCE: 210121.427C21  
;; CURRENT APPLICATION NUMBER: US/09/685,166A  
;; CURRENT FILING DATE: 2000-10-10  
;; NUMBER OF SEQ ID NOS: 898  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 72  
;; LENGTH: 511  
;; TYPE: DNA  
;; ORGANISM: Homo sapien  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1)..(511)  
;; OTHER INFORMATION: n = A,T,C or G  
US-09-685-166A-72

Query Match 43.2%; Score 444; DB 3; Length 511;  
Best Local Similarity 91.3%; Pred. No. 1,7e-127;  
Matches 462; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 492 CTTGCTCTCAAGAGTGTGTCATATTTTCCATCATCAGCTGTCTGCTTAGAAGAAC 551  
DB 506 CNGGCTCTCAAGAGGGGGGTATTTTNGCCATCATCAGCTGNGTTGCTTAGAAGAACG 447  
QY 552 GGCCTTCTGCTGCAATGAGAGAAATCATATACAGAGGTCAGAGAGGCAATCTTTT 611  
DB 446 GGTTCATGCTGCAATGAGAGAAATCATATACAGAGGTCAGAGAGGCAATCTTTT 387  
QY 612 CCTCATGCTATTTGTCCTTGAAGAGCTCTTGAAGATCTAGTTGGGCTTTCTTCTGG 671  
DB 386 CTTTCATGCTATTTGTCCTTGAAGAGCTTCTGAAGATCTAGTTGGGCTTTCTTCTGG 327  
QY 672 GTTTGGGCAATTCAGTTCTCATGTGTGACTATTTATCATATTTATTAACGTTTTC 731  
DB 326 GTTTGGGCAATTCAGTTCTCATGTGTGACTATTTATCATATTTATTAACGTTTTC 267  
QY 732 AAACCAAGGGGACACAGAGAACTCTGTAATTAACAATGAGAAATAGCCAGGCGA 791  
DB 266 AAACCAAGGGGACACAGAGAACTCTGTAATTAACAATGAGAAATAGCCAGGCGA 207  
QY 792 TCTCCAGACCAATCTCTCATGTTTTCACAGAGCTCTCCAGCCAAATAGCGCCT 851  
DB 206 TCTCCAGACCAATCTCTCATGTTTTCACAGAGCTCTCCAGCCAAATAGCGCCT 147  
QY 852 GCATAGTGTAGACATCTCGGCTTCTAGGCTTGTCTCTTAAATGCTTTTAAATCA 911  
DB 146 GNTATAGTGTAGACATCTCGGCTTCTAGGCTTGTCTCTTAAATGCTTTTAAATCA 87  
QY 912 GATAACGCTGGAAGGCTTTCATTTTACAGGCTGGAAGAGCTCTTGTGCTAGTTGA 971  
DB 86 GATAACGCTGGAAGGCTTTCATTTTACAGGCTGGAAGAGCTTGTGCTAGTTGA 27  
QY 972 ATTATGAGTGTGTTTCCGTAAATA 997  
DB 26 ATTATGAGTGTGTTTCCGTAAATA 1

RESULT 11  
US-09-115-453-72/c  
; Sequence 72, Application US/09115453B  
; Patent No. 6657056  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, David C.  
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND  
; TITLE OF INVENTION: METHODS FOR THEIR USE  
; FILE REFERENCE: 210121.427C4  
; CURRENT APPLICATION NUMBER: US/09/115,453B

;; CURRENT FILING DATE: 1998-07-14  
;; NUMBER OF SEQ ID NOS: 228  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 72  
;; LENGTH: 511  
;; TYPE: DNA  
;; ORGANISM: Homo sapien  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1)...(511)  
;; OTHER INFORMATION: n = A,T,C or G  
US-09-115-453-72

Query Match 43.2%; Score 444; DB 3; Length 511;  
Best Local Similarity 91.3%; Pred. No. 1,7e-127;  
Matches 462; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 492 CCTGCTCTCAAGAGTGTGTCTATTTTGGCATCATCCTGTCTGTCTTGAAGAAC 551  
DB 506 CNGGCTCTCAAGAGGAGGAGTATTTTNGCATCATCCTGNGTGTGTTAGAGAACG 447  
QY 552 GGGCTTCTGCTGCAATGAGAGAAATCATTAACAGAGCGTGACACAGAGGCCATCTTTT 611  
DB 446 GGTTCCTGCTGCAATGAGAGAAATCATTAACAGAGCGTGACACAGAGGCCATCTTTT 387  
QY 612 CCTCATCGTTATTTGCTCCCTAGAGCGTCTTGAAGATCTAGTTGGCTTTCTTTCTGG 671  
DB 386 CTTCATCGTTATTTGCTCCCTAGAGCGTCTTGAAGATCTAGTTGGCTTTCTTTCTGG 327  
QY 672 GTTTGGGCAATTCAGTTCTCATGTGTGTATTTATTTATTAATTAACGCTTTTC 731  
DB 326 GTTTGGGCAATTCAGTTCTCATGTGTGTATTTATTTATTAATTAACGCTTTTC 267  
QY 732 AAACAGTGGGACACAGAGAACTCTGTAATTAACAATAGGAAATAGCCAGCGCA 791  
DB 266 AAACAGTGGGACACAGAGAACTCTGTAATTAACAATAGGAAATAGCCAGCGCA 207  
QY 792 TCTCCAGACCAATCTCTCATGTGTGTATTTTCCACAGCTCTCCAGCAACCAATAGCGCT 851  
DB 206 TCTCCAGACCAATCTCTCATGTGTGTATTTTCCACAGCTCTCCAGCAACCAATAGCGCT 147  
QY 852 GCTATAGTGAACATCTCGGCTTCTAGCTTCTCTCTTCTAGTGTCTTAATCA 911  
DB 146 GNTATAGTGAACATCTCGGCTTCTAGCTTCTCTCTTCTAGTGTCTTAATCA 87  
QY 912 GATPACTGCTGGAAGCTTTTATTTTACAGCGCTGAGAGAGTCTTTCTAGTTGA 971  
DB 86 GATPACTGCTGGAAGCTTTTATTTTACAGCGCTGAGAGAGTCTTTCTAGTTGA 27  
QY 972 ATTATGCTGT 997  
DB 26 ATTATGCTGT 1

RESULT 12  
US-09-688-489-72/c  
;; Sequence 72, Application US/09688489  
;; Patent No. 6664377  
;; GENERAL INFORMATION:  
;; APPLICANT: Xu, Jiangchun  
;; APPLICANT: Dillon, Davin C.  
;; APPLICANT: Mitcham, Jennifer Lynn  
;; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE  
;; FILE REFERENCE: 210121.427D2  
;; CURRENT APPLICATION NUMBER: US/09/688,489  
;; CURRENT FILING DATE: 2000-10-13  
;; NUMBER OF SEQ ID NOS: 338  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 72  
;; LENGTH: 511  
;; TYPE: DNA  
;; ORGANISM: Homo sapien

;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1)...(511)  
;; OTHER INFORMATION: n = A,T,C or G  
US-09-688-489-72

Query Match 43.2%; Score 444; DB 3; Length 511;  
Best Local Similarity 91.3%; Pred. No. 1,7e-127;  
Matches 462; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 492 CCTGCTCTCAAGAGTGTGTCTATTTTGGCATCATCCTGTCTGTCTTGAAGAAC 551  
DB 506 CNGGCTCTCAAGAGGAGGAGTATTTTNGCATCATCCTGNGTGTGTTAGAGAACG 447  
QY 552 GGGCTTCTGCTGCAATGAGAGAAATCATTAACAGAGCGTGACACAGAGGCCATCTTTT 611  
DB 446 GGTTCCTGCTGCAATGAGAGAAATCATTAACAGAGCGTGACACAGAGGCCATCTTTT 387  
QY 612 CCTCATCGTTATTTGCTCCCTAGAGCGTCTTGAAGATCTAGTTGGCTTTCTTTCTGG 671  
DB 386 CTTCATCGTTATTTGCTCCCTAGAGCGTCTTGAAGATCTAGTTGGCTTTCTTTCTGG 327  
QY 672 GTTTGGGCAATTCAGTTCTCATGTGTGTATTTATTTATTAATTAACGCTTTTC 731  
DB 326 GTTTGGGCAATTCAGTTCTCATGTGTGTATTTATTTATTAATTAACGCTTTTC 267  
QY 732 AAACAGTGGGACACAGAGAACTCTGTAATTAACAATAGGAAATAGCCAGCGCA 791  
DB 266 AAACAGTGGGACACAGAGAACTCTGTAATTAACAATAGGAAATAGCCAGCGCA 207  
QY 792 TCTCCAGACCAATCTCTCATGTGTGTATTTTCCACAGCTCTCCAGCAACCAATAGCGCT 851  
DB 206 TCTCCAGACCAATCTCTCATGTGTGTATTTTCCACAGCTCTCCAGCAACCAATAGCGCT 147  
QY 852 GCTATAGTGAACATCTCGGCTTCTAGCTTCTCTCTTCTAGTGTCTTAATCA 911  
DB 146 GNTATAGTGAACATCTCGGCTTCTAGCTTCTCTCTTCTAGTGTCTTAATCA 87  
QY 912 GATPACTGCTGGAAGCTTTTATTTTACAGCGCTGAGAGAGTCTTTCTAGTTGA 971  
DB 86 GATPACTGCTGGAAGCTTTTATTTTACAGCGCTGAGAGAGTCTTTCTAGTTGA 27  
QY 972 ATTATGCTGT 997  
DB 26 ATTATGCTGT 1

RESULT 13  
US-09-679-426-72/c  
;; Sequence 72, Application US/09679426  
;; Patent No. 679515  
;; GENERAL INFORMATION:  
;; APPLICANT: Xu, Jiangchun  
;; APPLICANT: Dillon, Davin C.  
;; APPLICANT: Mitcham, Jennifer L.  
;; APPLICANT: Harlocker, Susan L.  
;; APPLICANT: Jiang, Yugu  
;; APPLICANT: Henderson, Robert A.  
;; APPLICANT: Kalos, Michael D.  
;; APPLICANT: Fanger, Gary R.  
;; APPLICANT: Retter, Marc W.  
;; APPLICANT: Stolk, John A.  
;; APPLICANT: Day, Craig H.  
;; APPLICANT: Vedvik, Thomas S.  
;; APPLICANT: Carter, Darrick  
;; APPLICANT: Li, Samuel  
;; APPLICANT: Wang, Aijun  
;; APPLICANT: Skeiky, Yasir A.W.  
;; APPLICANT: Hepler, William  
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
;; FILE REFERENCE: 210121.427C20  
;; CURRENT APPLICATION NUMBER: US/09/679,426

CURRENT FILING DATE: 2000-10-02  
NUMBER OF SEQ ID NOS: 895  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 72  
LENGTH: 511  
TYPE: DNA  
ORGANISM: Homo sapien  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1) - (511)  
OTHER INFORMATION: n = A,T,C or G  
US-09-679-426-72

Query Match 43.2%; Score 444; DB 3; Length 511;  
Best Local Similarity 91.3%; Pred. No. 1.7e-127;  
Matches 462; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 492 CCTGCTCTCAAGAGTGTCTATTTTGGCATCATCACTGCTGTCTGCTTAGAAGAC 551  
Db 506 CNGGTCCTTCAAGAGGGGTGTATTTTNGCATCATCACTGCTGTCTTAGAAGACG 447  
Qy 552 GCGTTTCTGCTGCAATGAGAGAAATCATACAGCGGTGGCAAGAGGCCATCTTTT 611  
Db 446 GGTATTNGCTGCAATGAGAGAAATCATACAGCGGTGGCAAGAGGCCATCTTTT 387  
Qy 612 CCTCATCGTATTATGTCCTTGAAGCGCTCTTCTGAGATCTAGTTGGCTTTCTTCTG 671  
Db 386 CTTATCGGTATTATGTCCTTGAAGCGCTTCTGAGATCTAGTTGGCTTTCTTCTG 327  
Qy 672 GTTTGGGCATTTTCAGTTCTCATGTGTGTACTATTCTATTCATTATTTGATAACGGTTTC 731  
Db 326 GTTTGGGCATTTTCAGTTCTCATGTGTGTACTATTCTATTCATTATTTGATAACGGTTTC 267  
Qy 732 AAACAGTGGGCAACAGAGAACTCTGTATTAACAATGAGAAATAGCCAGCGCA 791  
Db 266 AAACAGTGGGCAACAGAGAACTCTGTATTAACAATGAGAAATAGCCAGCGCA 207  
Qy 792 TCTCCAGACCAATCTCTCATGTTTTCACAGCTCTCCAGCCAAACCAATAGCGCT 851  
Db 206 TTTCCAGACCAATCTCTCATGTTTTCACAGCTCTCCAGCCAAACCAATAGCGCT 147  
Qy 852 GCTATAGTGAACATCTCGGGCTTTCAGCTTGTCTCTCTTTAGTGTCTTAAATCA 911  
Db 146 GNTATAGTGAACATCTCGGGCTTTCAGCTTGTCTCTCTTTAGTGTCTTAAATCA 87  
Qy 912 GATACTGCTGGAAGCCCTTTCATTTTACAGCCCTGGAAGAGTCTTCTTGTAGTTGA 971  
Db 86 GATACTGCTGGAAGCCCTTTCATTTTACAGCCCTGGAAGAGTCTTCTTGTAGTTGA 27  
Qy 972 ATTATGTGTGTGTTTTCCGTAATA 997  
Db 26 ATTATGTGTGTGTTTTCCGTAATA 1

RESULT 14  
US-09-759-143-72/c  
Sequence 72, Application US/09759143  
Patent No. 6800746  
GENERAL INFORMATION:

APPLICANT: Xu, Jianshun  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Jiansun, Yuqi  
APPLICANT: Henderson, Robert A.  
APPLICANT: Kalos, Michael D.  
APPLICANT: Fanger, Gary R.  
APPLICANT: Retter, Marc W.  
APPLICANT: Stolk, John A.  
APPLICANT: Day, Craig H.  
APPLICANT: Vedvik, Thomas S.  
APPLICANT: Carter, Darrick  
APPLICANT: Li, Samuel

APPLICANT: Wang, Aijun  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Hepler, William  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.427C23  
CURRENT APPLICATION NUMBER: US/09/759.143  
CURRENT FILING DATE: 2001-01-12  
NUMBER OF SEQ ID NOS: 934  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 72  
LENGTH: 511  
TYPE: DNA  
ORGANISM: Homo sapien  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1) - (511)  
OTHER INFORMATION: n = A,T,C or G  
US-09-759-143-72

Query Match 43.2%; Score 444; DB 3; Length 511;  
Best Local Similarity 91.3%; Pred. No. 1.7e-127;  
Matches 462; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 492 CCTGCTCTCAAGAGTGTCTATTTTGGCATCATCACTGCTGTCTGCTTAGAAGAC 551  
Db 506 CNGGTCCTTCAAGAGGGGTGTATTTTNGCATCATCACTGCTGTCTTAGAAGACG 447  
Qy 552 GCGTTTCTGCTGCAATGAGAGAAATCATACAGCGGTGGCAAGAGGCCATCTTTT 611  
Db 446 GGTATTNGCTGCAATGAGAGAAATCATACAGCGGTGGCAAGAGGCCATCTTTT 387  
Qy 612 CCTCATCGTATTATGTCCTTGAAGCGCTCTTCTGAGATCTAGTTGGCTTTCTTCTG 671  
Db 386 CTTATCGGTATTATGTCCTTGAAGCGCTTCTGAGATCTAGTTGGCTTTCTTCTG 327  
Qy 672 GTTTGGGCATTTTCAGTTCTCATGTGTGTACTATTCTATTCATTATTTGATAACGGTTTC 731  
Db 326 GTTTGGGCATTTTCAGTTCTCATGTGTGTACTATTCTATTCATTATTTGATAACGGTTTC 267  
Qy 732 AAACAGTGGGCAACAGAGAACTCTGTATTAACAATGAGAAATAGCCAGCGCA 791  
Db 266 AAACAGTGGGCAACAGAGAACTCTGTATTAACAATGAGAAATAGCCAGCGCA 207  
Qy 792 TCTCCAGACCAATCTCTCATGTTTTCACAGCTCTCCAGCCAAACCAATAGCGCT 851  
Db 206 TTTCCAGACCAATCTCTCATGTTTTCACAGCTCTCCAGCCAAACCAATAGCGCT 147  
Qy 852 GCTATAGTGAACATCTCGGGCTTTCAGCTTGTCTCTCTTTAGTGTCTTAAATCA 911  
Db 146 GNTATAGTGAACATCTCGGGCTTTCAGCTTGTCTCTCTTTAGTGTCTTAAATCA 87  
Qy 912 GATACTGCTGGAAGCCCTTTCATTTTACAGCCCTGGAAGAGTCTTCTTGTAGTTGA 971  
Db 86 GATACTGCTGGAAGCCCTTTCATTTTACAGCCCTGGAAGAGTCTTCTTGTAGTTGA 27  
Qy 972 ATTATGTGTGTGTTTTCCGTAATA 997  
Db 26 ATTATGTGTGTGTTTTCCGTAATA 1

RESULT 15  
US-09-651-236-72/c  
Sequence 72, Application US/09651236  
Patent No. 6818751  
GENERAL INFORMATION:

APPLICANT: Xu, Jianshun  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Jiansun, Yuqi  
APPLICANT: Henderson, Robert A.  
APPLICANT: Kalos, Michael D.

```
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.42718C18
; CURRENT APPLICATION NUMBER: US/09/651.236
; CURRENT FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 865
; SOFTWARE: FaalSEQ for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 511
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(511)
; OTHER INFORMATION: n = A,T,C or G
US-09-651-236-72
```

Query Match 43.2%; Score 444; DB 3; Length 511;

Best Local Similarity 91.3%; Pred. No. 1.7e-127; Matches 462; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

```
QY 492 CCGCTCTCTCAAGAGTGTGCTATTTTGGCATCATCACTGCTGTCTGCTTAAGAAC 551
Db 506 CNGGTNCTTCAAGAGGGGGGTATTTTNGCATCATCACTGNGTTGCTTAGAAGACG 447
QY 552 GCGTTTCTGCTGATGAGAGAAATCATTAACAGCGGTGGCACAAGAGGCCATCTTTT 611
Db 446 GGTTTNNGCTGCAATGAGAGAAATCATTAACAGCGGGGCAAGAGGCCATCTTTT 387
QY 612 CCTCATGGTTATTTGCTCCCTTAAGAGCGTCTTCTAGAGATCTAGTGGGCTTTCTTCTGG 671
Db 386 CTTCATCGGTATTTGCTCCCTTAAGAGCGTCTTCTAGAGATCTAGTGGGCTTTCTTCTGG 327
QY 672 GTTGGGCACTTCAAGTTCATGAGTGTACTATTTCTATCATTTATGATTAAGGTTTC 731
Db 326 GTTGGGCACTTCAAGTTCATGAGTGTACTATTTCTATCATTTATGAGTGTTC 267
QY 732 AAACCAAGTGGGCAACAGAACTCACTGTATTAACAATGAGATAGCCAGGCGA 791
Db 266 AAACCAAGTGGGCAACAGAACTCACTGTATTAACAATGAGATAGCCAGGCGA 207
QY 792 TTTCCAGCACCAATCTCTCCATGTTTTCCACAGGCTCTCCAGCCCAACCAATAAGCGCT 851
Db 206 TTTCCAGCACCAATCTCTCCATGTTTTCCACAGGCTCTCCAGCCCAACCAATAAGCGCT 147
QY 852 GCTATAGTGAATCTCTGCGGCTTCAAGCTTGTCCCTCTTAAAGTGTCTTAATCA 911
Db 146 GNTATAGTGAATCTCTGCGGCTTCAAGCTTGTCCCTCTTAAAGTGTCTTAATCA 87
QY 912 GATAAGTGGTGAAGCTTCAATTTTACAGCCCTGAAGAGAGCTTCTTGTAGTTGA 971
Db 86 GATAAGTGGTGAAGCTTCAATTTTACAGCCCTGAAGAGAGCTTCTTGTAGTTGA 27
QY 972 ATTATGTGTGTGTTTTCCGTAATA 997
Db 26 ATTATGTGTGTGTTTTCCGTAATA 1
```

Search completed: December 10, 2005, 23:27:03  
Job time : 232 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2005, 20:04:37 ; Search time 4078 Seconds  
(Without alignments)  
11782.818 Million cell updates/sec

Title: US-10-031-158B-13

Perfect score: 1027  
Sequence: 1 gggcagaagctgggcacaaa.....attcaaaaaatgaaaagct 1027

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: gb\_esc1:\*  
2: gb\_esc2:\*  
3: gb\_esc3:\*  
4: gb\_esc4:\*  
5: gb\_esc5:\*  
6: gb\_esc6:\*  
7: gb\_esc7:\*  
8: gb\_esc8:\*  
9: gb\_esc9:\*  
10: gb\_esc10:\*  
11: gb\_esc11:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	# Query	Match Length	DB ID	Description
1	961.8	93.7	3533	4	BC030554 Homo sapi
2	880.8	85.8	1510	4	BC017861 Homo sapi
3	837.8	81.6	914	5	BO934484 AGENCOURT
4	636	61.9	821	2	BF679123
5	627.8	60.1	997	2	BF678971
6	618.6	60.2	636	1	A1972955
7	614	59.8	938	1	BF674593
8	613.8	59.8	938	1	A1685999
9	605.2	58.9	959	7	CN645427
10	590.2	57.5	703	2	BC217853
11	589.6	57.4	629	1	A1768834
12	584	56.9	700	6	CA449324
13	580.4	56.5	582	2	BF326754
14	578.8	56.4	613	3	BI905966
15	569.8	55.5	719	2	BF681238
16	567.8	55.3	670	2	BF677915
17	565.6	55.1	789	5	BU199143
18	561	54.6	629	2	BF679165
19	558.4	54.4	1058	3	BM544213
20	553	53.8	849	2	BF677648
21	552.6	53.8	849	2	BF679284
22	548.8	53.4	803	2	BF680840

c	23	547.6	53.3	701	10	AG175832	Pan tlog1
	24	543.6	52.9	857	2	BF674457	BF674457 602137231
	25	540.2	52.6	861	2	BF681385	BF681385 602156579
	26	532.4	51.8	665	1	AA569813	AA569813 nm46c02.8
	27	530.8	51.7	600	1	AI244217	AI244217 q186c05.x
	28	516	50.2	1055	2	BF664904	BF664904 602268596
	29	493.4	48.0	927	2	BF679476	BF679476 602153280
	30	490.8	47.8	572	2	BF370026	BF370026 MK3-FN00
	31	490.4	47.8	503	1	AI253507	AI253507 ac31c11.x
	32	490	47.7	525	5	EX112931	EX112931 BX112931
	33	488.4	47.6	513	1	AI823897	AI823897 wJ28e10.x
	34	466	45.4	479	1	AM575113	AM575113 UI-HF-BKO
	35	449.4	43.8	596	1	AV715641	AV715641 AV715641
	36	448.6	43.7	529	1	AA908306	AA908306 og33h07.s
	37	435.6	42.4	442	1	AI718234	AI718234 ad42h05.x
	38	434.8	42.3	497	1	AA654656	AA654656 nc76c02.8
	39	429.2	41.8	512	1	AA657507	AA657507 nc64g07.8
	40	428.6	41.7	681	1	AV714756	AV714756 AV714756
	41	428.4	41.7	870	2	BF678129	BF678129 602085181
	42	428	41.7	460	2	BE772783	BE772783 RCI-PT013
	43	420.4	40.9	985	1	AI557112	AI557112 pr2.1.13
	44	418.6	40.8	523	2	BF675605	BF675605 602083495
	45	416.8	40.6	457	8	N50880	N50880 yy92d06.s1

#### ALIGNMENTS

RESULT 1	BC030554	3533 bp	mRNA	linear	HTC 20-MAY-2002
LOCUS	BC030554				
DEFINITION	Homo sapiens, similar to T cell receptor gamma locus, clone IMAGE:5213435, mRNA.				
ACCESSION	BC030554.1	GI:20988582			
VERSION	BC030554.1				
KEYWORDS	HTC.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 3533)				
AUTHORS	Strausberg, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (07-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov				
COMMENT	Contact: MGC help desk Email: cgabs-remail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: http://www.nisc.nih.gov/ Contact: nisc.mgc@hgrl.nih.gov Ahter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Bakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Leric, P., Legaepi, R., Maduro, O.L., Maitello, C., Maskeri, B., Mastrisan, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantirpop, S., Thomas, P.J., Touchman, J.W., Tsunggeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Series: IRAX Plate: 64 Row: k Column: 18 This clone has the following problem: frame shifted. Location/Qualifiers				
FEATURES					

Source	1..3533	organism="Homo sapiens"	
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5233435"		
	/tissue_type="Blood, adult leukocytes"		
	/clone_id="NH_MGC_118"		
	/lab_host="DH10B"		
	/note="Vector: pCMV-SPORT6"		
ORIGIN			
Query Match	93.7%	Score 961.8; DB 4; Length 3533;	
Best Local Similarity	99.4%	Pred. NO. 7.9e-252;	
Matches	963; Conservative	0; Mismatches 6; Indels 0; Gaps 0;	
59	CAGATAAACA	CTGATGCAAGATGTTTCCCCAACCCACATATTTTCTTCCTCAATTG 118	
Db	743	CTGATAAACA	CTTATGATGACATGTTTCCCCAACCCACATATTTTCTTCCTCAATTG 802
Qy	119	CTGAACAAG	AGCTCCAGAAGCTGGAAACATACCTTTGTCTTCTTGAGAAATTTTTCCTG 178
Db	803	CTGAACAAG	AGCTCCAGAAGCTGGAAACATACCTTTGTCTTCTTGAGAAATTTTTCCTG 862
Qy	179	ATGTTATTA	AGATTAAGTATCATTTGGCAAGAAAGAGAGCAACGATTTGGGATCCAGAGG 238
Db	863	ATGTTATTA	AGATTAAGTATCATTTGGGAGAAAGAGAGCAACGATTTGGGATCCAGAGG 922
Qy	239	GGAACACCA	TGAGAGACTAAGCAACATACATGAAATTTAGCTGTTAACGTCGCAGAA 298
Db	923	GGAACACCA	TGAGAGACTAAGCAACATACATGAAATTTAGCTGTTAACGTCGCAGAA 982
Qy	289	AGTCACTGA	CAAGAAACAACAGATGTATCGTCAGACATGAGATATATTAACCGAGATTG 358
Db	983	AGTCACTGA	CAAGAAACAACAGATGTATCGTCAGACATGAGATATATTAACCGAGATTG 1042
Qy	359	ATCAAGAA	AAATTAATCTTCTCTCCAAATTAAGACGAGTGCATCAACATGATCCCAAGCA 418
Db	1043	ATCAAGAA	AAATTAATCTTCTCTCCAAATTAAGACAGATGTATCAACATGATCCCAAGCA 1102
Qy	419	ATTGTTCAA	AAAGATCAAAATGATATACATCTGCTGACGTCAACAACCTCTGCATATT 478
Db	1103	ATTGTTCAA	AAAGATCAAAATGATATACATCTGCTGACGTCAACAACCTCTGCATATT 1162
Qy	479	ACATGTACC	TCTCTCTGCTCTCAAGAGTGTGCTATTTTGGCATCATCCTGCTGTC 538
Db	1163	ACATGTACC	TCTCTCTGCTCTCTCAAGAGTGTGCTATTTTGGCATCATCCTGCTGTC 1222
Qy	539	TGCTTAA	GAAAGACGGCTTTCTGCTGCATATGAGAGAAATCAATAAGACGGTGGACAAG 598
Db	1223	TGCTTAA	GAAAGACGGCTTTCTGCTGCATATGAGAGAAATCAATAAGACGGTGGACAAG 1282
Qy	599	GAGGCAAT	CTTTTCTCTCATCGGTTATTGTCCTTAAGAGCTTTCTGAGATCTAGTTGG 658
Db	1283	GAGGCAAT	CTTTTCTCTCATCGGTTATTGTCCTTAAGAGCTTTCTGAGATCTAGTTGG 1344
Qy	659	GCTTCTTT	CTTGAGGTTTGGGCATTTCAGTTTCTCATGTGTGACTATTTCTATCATTTG 718
Db	1343	GCTTCTTT	CTTGAGGTTTGGGCATTTCAGTTTCTCATGTGTGACTATTTCTATCATTTG 1402
Qy	719	TATTAACG	TTTTCAACACAGTGGGACACAGAAACCTGACTCTGTATTAACATGAGGA 778
Db	1403	TATTAACG	TTTTCAACACAGTGGGACACAGAAACCTGACTCTGTATTAACATGAGGA 1462
Qy	779	ATAGCAAG	GGGATCTCCAGACAACAATCTCTCAATGTTTCCACAGCTCTCCAGCAAC 838
Db	1463	ATAGCAAG	GGGATCTCCAGACAACAATCTCTCAATGTTTCCACAGCTCTCCAGCAAC 1522
Qy	839	CCAAATAG	CGCTGCTATATGTTAGACATCTGCGGCTTCTAGCTTGTCCCTCTCTTAG 898
Db	1523	CCAAATAG	CGCTGCTATATGTTAGACATCTGCGGCTTCTAGCTTGTCCCTCTCTTAG 1582
Qy	899	TGTTCTTA	TATCAGATTAATGCTGCTGGAGCTTTCAATTTTACACGCTCCTGAAGACGCTT 958

Db	1583	TGTTCTTTAATTCAGATTAAGTCGCGAAGCCTTTCATTTTAACAAGCCCTCAAGACGCTT	16412
Oy	959	CTTTGCTAGATTGGAATTATGTGCTGTCTTTTCCGTATTAAGCAAAATTAATTTAAAAAA	1018
Db	1643	CTTTGCTAGATTGGAATTATGTGCTGTCTTTTCCGTATTAAGCAAAATTAATTTAAAAAA	1702
Oy	1019	TGAAAAGTT 1027	
Db	1703	TGAAAAGTT 1711	
RESULT 2			
BC017861			
LOCUS			
DEFINITION	Homo sapiens, similar to T cell receptor gamma locus, clone	1510 bp	mRNA
IMAGE:4687960, mRNA.			
ACCESSION	BC017861		
VERSION	BC017861.1	GI:17389678	
KEYWORDS	HTC.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Hominidae; Homo.		
REFERENCE	1 (bases 1 to 1510)		
AUTHORS	Strausberg,R		
TITLE	Direct Submission		
JOURNAL	Submitted (03-DEC-2001) National Institutes of Health, Mammalian		
	Gene Collection (MGC), Cancer Genomics Office, National Cancer		
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,		
	USA		
REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>		
COMMENT	Contact: MGC help desk		
	Email: <a href="mailto:cgaps-remail.nih.gov">cgaps-remail.nih.gov</a>		
	Tissue Procurement: CLONTECH		
	cDNA Library Preparation: CLONTECH Laboratories, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)		
	DNA Sequencing by: Sequencing Group at the Stanford Human Genome		
	Center, Stanford University School of Medicine, Stanford, CA 94305		
	Web site: <a href="http://www-shgc.stanford.edu">http://www-shgc.stanford.edu</a>		
	Contact: (Dickson, Mark) <a href="mailto:mcd@paxil.stanford.edu">mcd@paxil.stanford.edu</a>		
	Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,		
	R. M.		
FEATURES			
source	Clone distribution: MGC clone distribution information can be found		
	through the I.M.A.G.E. Consortium/LINL at: <a href="http://image.lnl.gov">http://image.lnl.gov</a>		
	Series: IRAL Plate: 36 Row: k Column: 19		
	This clone has the following problem: frame shifted.		
	Location/Qualifiers		
	1. .1510		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:4687960"		
	/tissue_type="Lung"		
	/clone_id="NIH_MGC_77"		
	/lab_host="DH10B"		
	/note="Vector: pDNR-LIB"		
ORIGIN			
Query Match	85.8%; Score 880.8; DB 4; Length 1510;		
Best Local Similarity	93.6%; Pred. No. 1e-229;		
Matches 949; Conservative	0; Mismatches 17; Indels 48; Gaps 1;		
Oy	59	CAGATAAACAATTGATGCAAGATGTTTCCCAAGCCACATATTTCTTCCTCAATG 118	
Db	464	CTGATTAACAACCTTGATGCAAGATGTTTCCCAAGCCACATATTTCTTCCTCAATG 523	
Oy	119	CTGAAAACAAAGCTCAGAAAGCTGGAACATACCTTTGCTCTTGAGAAATTTTCCCTG 178	
Db	524	CTGAAAACAAACCTCAGAAAGCTGGAACATACCTTTGCTCTTGAGAAATTTTCCAG 583	
Oy	179	ATGTTATTAGATCATTGGCAAGAAAAGAGCAACACGATCTTCGGATCCAGAGAG 238	

```

Db      584 ATATTATTAAAGATATGATGGCAAGAAAGAGCAACAGATTCCTGGATCCAGAGAGG 643
Qy      239 GGAACACCATGAAAGACTTAACGACACATCACTGAAATTTAGCTGTAAACGCTGCAGAAA 298
Db      644 GGAACACCATGAAAGACTTAACGACACATCACTGAAATTTAGCTGTAAACGCTGCAGAAA 703
Qy      299 AGTACATGGAACAAGAACACAGATGTATCGACACATGAAATTAATTAAGACGAGTTG 358
Db      704 AGTACATGGAACAAGAACACAGATGTATCGACACATGAAATTAATTAAGACGAGTTG 763
Qy      359 ATCAAGAAATTAATCTTTCTCTCAATAAAGAC----- 389
Db      764 ATCAAGAAATTAATCTTTCTCTCAATAAAGACATGTACACACAGTGGATCCAAATACA 823
Qy      390 -----GGATGTCAATCAATGATGATCCCAAGACAAATTTGTCAAAAG 430
Db      824 ATTAATCAAGATGCAAAATGATGTCAATGATGATCCCAAGACAAATTTGTCAAAAG 883
Qy      431 ATGCAATGATGATCACTGCTGAGCTCAAAACACCTGCTGATTAAGATGATGCTCC 490
Db      884 ATGCAATGATGATCACTGCTGAGCTCAAAACACCTGCTGATTAAGATGATGATGCTCC 943
Qy      491 TCCTGCTCTCAAGAGTGTGTCTATTGTCATCATCACTGCTGCTGCTTGAAGAA 550
Db      944 TCCTGCTCTCAAGAGTGTGTCTATTGTCATCATCACTGCTGCTGCTTGAAGAA 1003
Qy      551 CGGCTTTCTGCTGCAATGAGAGAAATCATPAACAGAGTGGCAAGAGGCCATCTTT 610
Db      1004 CGGCTTTCTGCTGCAATGAGAGAAATCATPAACAGAGTGGCAAGAGGCCATCTTT 1063
Qy      611 TCCTCATCGGTATTTGTCCTGAGAAAGCTCTTGAAGATCTAATGAGCTTTCTTTCTG 670
Db      1064 TCCTCATCGGTATTTGTCCTGAGAAAGCTCTTGAAGATCTAATGAGCTTTCTTTCTG 1123
Qy      671 GGTGGGCAATTCAGTCTCATGTGTGATCAATTCATCAATTAATGATTAACGTTT 730
Db      1124 GGTGGGCAATTCAGTCTCATGTGTGATCAATTCATCAATTAATGATTAACGTTT 1183
Qy      731 CAACACAGTGGGCAACAGAGAACTCTGTATTAACAGAGAAATGACAGGCG 790
Db      1184 CAACACAGTGGGCAACAGAGAACTCTGTATTAACAGAGAAATGACAGGCG 1243
Qy      791 ATTCACGACCAATCTCTCATGTGTTTCCACAGCTCTCCAGCAACCCAAATAGGCC 850
Db      1244 ATTCACGACCAATCTCTCATGTGTTTCCACAGCTCTCCAGCAACCCAAATAGGCC 1303
Qy      851 TGCTATGTGTGACATCTGCGGCTCTAGCTGTGCTGCTCTGTTAGTGTCTTTATC 910
Db      1304 TGCTATGTGTGACATCTGCGGCTCTAGCTGTGCTGCTCTGTTAGTGTCTTTATC 1363
Qy      911 AGATTAATGCTGGAAGCTTTCAATTTTACAGGCTGTAAGACAGCTTTTGTAGTTG 970
Db      1364 AGATTAATGCTGGAAGCTTTCAATTTTACAGGCTGTAAGACAGCTTTTGTAGTTG 1423
Qy      971 AATTATGTGTGTGTTTTCCTTAATTAAGCAAAATTAATTAATTAATTAATTAATTA 1024
Db      1424 AATTATGTGTGTGTTTTCCTTAATTAAGCAAAATTAATTAATTAATTAATTAATTA 1477

```

```

REFERENCE 1 (bases 1 to 914)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-rc@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLM13628 row: c column: 11
High quality sequence stop: 633.
FEATURES
Source
location/Qualifiers
1..914
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6205378"
/sex="male"
/tissue_type="sciatic nerve"
/dev_stage="adult, 70 yr"
/lab_host="DH10B"
/clone_lib="Lupski_sciatic nerve"
/notes="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACGCGTCCG-3' and
5'-GACTGATCTTCAATGCGAGCGGCGCCCTT(15)-3'. Size selected >
1 kb for average insert length 1.87 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."
ORIGIN
Query Match 81.6%; Score 837.8; DB 5; Length 914;
Best Local Similarity 98.5%; Pred. No. 5,4e-218;
Matches 856; Conservative 0; Mismatches 12; Indels 1; Gaps 1;
Qy 157 CTTCTTGAGAAATTTTTCCTGATGTTATTAAGATACATTGGCAAGAAAGAGCAAC 216
Db 1 CTTCTTGAGAAATTTTTCCTGATGTTATTAAGATACATTGGCAAGAAAGAGCAAC 60
Qy 217 ACGATTCGGGATCCAGAGGGGAAACCATGAAGACTTAAGCAACATACATGAATTT 276
Db 61 ACGATTCGGGATCCAGAGGGGAAACCATGAAGACTTAAGCAACATACATGAATTT 120
Qy 277 AGCTGTTAAGCGTGCCAGAAAGTCACTGGAACAAAGACAGATGTATCGTCAAGCAT 336
Db 121 AGCTGTTAAGCGTGCCAGAAAGTCACTGGAACAAAGACAGATGTATCGTCAAGCAT 180
Qy 337 GAGATTAATTAAGCGAGTGTATCAAGAAATATCTTCTCCTCAATTAAGAGCATGTC 396
Db 181 GAGATTAATTAAGCGAGTGTATCAAGAAATATCTTCTCCTCAATTAAGAGCATGTC 240
Qy 397 ATCAATGATGCCAAAGCAATTTGTTCAAAAGATGCAATGATACATGCTGTCAG 456
Db 241 ATCAATGATGCCAAAGCAATTTGTTCAAAAGATGCAATGATACATGCTGTCAG 300
Qy 457 CTGCAAAACCTCTGATATTACATGTAACCTCTCTGCTCTCAAGATGTGTTAT 516
Db 301 CTGCAAAACCTCTGATATTACATGTAACCTCTCTGCTCTCAAGATGTGTTAT 360
Qy 517 TTTGGCATCATCACTGCTGCTGCTTGAAGAAAGCGCTTCTGTCGCAATGAGAGAAA 576
Db 361 TTTGGCATCATCACTGCTGCTGCTTGAAGAAAGCGCTTCTGTCGCAATGAGAGAAA 420
Qy 577 TCATTAACAGCGGTGGCAAGAGAGCCATCTTTCTCTCATCGGTATTTGCTCTAGAG 636
Db 421 TCATTAACAGCGGTGGCAAGAGAGCCATCTTTCTCTCATCGGTATTTGCTCTAGAG 480

```

```

RESULT 3
LOCUS B0934484 914 bp mRNA linear EST 21-AUG-2002
DEFINITION AGNCOURT 8754342 Lupski_sciatic_nerve Homo sapiens cDNA clone
IMAGE:6205378 5', mRNA sequence.
ACCESSION B0934484
VERSION B0934484.1 GI:22349867
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

```

QY 637 CGTCTTGAGAGATCTAGTGGGCTTTCTTTCTGGGTTTGAGCCATTTCAGTTCATGT 696  
 DB 481 CGTCTTGAGAGATCTAGTGGGCTTTCTTTCTGGGTTTGAGCCATTTCAGTTCATGT 540  
 QY 697 GTGTACTATTTCTATCATTTATGTATTAAGGTTTTCAAACCGATGGGCACAGAGAACT 756  
 DB 541 GTGTACTATTTCTATCATTTATGTATTAAGGTTTTCAAACCGATGGGCACAGAGAACT 600  
 QY 757 CACTCTGTATTAACATAGAGAAATAGCCAGCGCATCTCCAGACCAATCTCTCCATGTT 816  
 DB 601 CACTCTGTATTAACATAGAGAAATAGCCAGCGCATCTCCAGACCAATCTCTCCATGTT 660  
 QY 817 TTCAACAGCTCTCCAGCCAAACCAATAGCGCTGTATAGTGAATCTGCGGCT 876  
 DB 661 TTCAACAGCTCTCCAGCCAAACCAATAGCGCTGTATAGTGAATCTGCGGCT 720  
 QY 877 TCTAGCCCTGTCCCTCTTGTAGTGTCTTAAATGAGTAATGAGTGAATCTGCGGCT 936  
 DB 721 TCTAGCCCTGTCCCTCTTGTAGTGTCTTAAATGAGTAATGAGTGAATCTGCGGCT 780  
 QY 937 TTACAGCGCCCTGAAGCAGTCTTCTTGTAGTGAATATGAGT-GTGTCTTCCGTAA 995  
 DB 781 TTACAGCGCCCTGAAGCAGTCTTCTTGTAGTGAATATGAGTGTGTGGGTTTTTCCGTAA 840  
 QY 996 TAAACAAATTAATTTAAAAAATGAAAA 1024  
 DB 841 TAAACAAATTAATTTAAAAAATGAAAA 869

RESULT 4  
 BF679123 821 bp mRNA linear EST 21-DEC-2000  
 LOCUS 602153390F1 NIH\_MGC\_83 Homo sapiens cDNA clone IMAGE:4294247 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BF679123  
 VERSION BF679123.1 GI:11953018  
 KEYWORDS EST.  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.  
 1 (bases 1 to 821)  
 NIH-MGC http://mgi.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 CONTACT: Robert Strausberg, Ph.D.  
 Email: rgs@dbp-remail.nih.gov  
 Tissue Procurement: CLONETECH Laboratories, Inc.  
 cDNA Library Preparation: CLONETECH Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMD)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/BLMD at:  
 http://image.llnl.gov  
 Plate: LMD1143 row: d column: 24  
 High quality sequence stop: 620.  
 Location/Qualifiers  
 1..821  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4294247"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NIH\_MGC\_83"  
 /note="Organ: prostate; Vector: pDNR-LIB (Clontech);  
 Site 1: SfiI (ggccgctcgcc); Site 2: SfiI  
 (ggccatcgcc); 5' and 3' adaptors were used in cloning  
 as follows: 5' adaptor sequence: 5'-CACGCGCATTTAGGCC-3'  
 and 3' adaptor sequence:  
 5'-ATTCTAGAGCCGAGCGCGCGCATG-dT(30)BN-3' (where B = A,  
 C, or G and N = A, C, G, or T). Average insert size 1.4

## ORIGIN

Query Match 61.9%; Score 636; DB 2; Length 821;  
 Best Local Similarity 95.4%; Pred. No. 1e-162;  
 Matches 731; Conservative 0; Mismatches 25; Indels 10; Gaps 7;

kb (range 0.5-4.0 kb). 14/15 colonies contained inserts  
 by PCR. This library was enriched for full-length clones  
 and was constructed by Clontech Laboratories (Palo Alto,  
 CA).

QY 12 GGGCAAAAATCAAGTATTTGGTCCCGAACAACCTTATCTACAGATAAACACT 71  
 DB 3 GGGCAAAAATCAAGTATTTGGTCCCGAACAACCTTATCTACAGATAAACACT 61  
 QY 72 TGATGAGATGTTTCCCGAAGCCCAATTTTCTTCTCAATTGCTGAAACAAAGCT 131  
 DB 62 TGATGAGATGTTTCCCGAAGCCCAATTTTCTTCTCAATTGCTGAAACAAAGCT 121  
 QY 132 CCAGAAAGCTGGAACATACCTTTGTCTTTGAGAAATTTTCCCTGATGTTAAGAT 191  
 DB 122 CCAGAAAGCTGGAACATACCTTTGTCTTTGAGAAATTTTCCCTGATGTTAAGAT 181  
 QY 192 ACATTGGCAAGAAAAGAAAGCAACAGATTCGGATCCAGAGGGGAAACACATGAA 251  
 DB 182 ACATTGGCAAGAAAAGAAAGCAACAGATTCGGATCCAGAGGGGAAACACATGAA 241  
 QY 252 GACTTAACGACATACATGAAATTTAGCTGTTAAACGTTGACAGAAAGTCACTGACA 311  
 DB 242 GACTTAACGACATACATGAAATTTAGCTGTTAAACGTTGACAGAAAGTCACTGACA 301  
 QY 312 AGAACACAGATGTTATGCTGACATGAGATTAATTAAGAGTTGATCAAGAAATAT 371  
 DB 302 AGAACACAGATGTTATGCTGACATGAGATTAATTAAGAGTTGATCAAGAAATAT 361  
 QY 372 CTTTCTCCCAATTAAGACGATGTCATCAACATGATCCCAAGACATTTGTTCAAAAGA 431  
 DB 362 CTTTCTCCCAATTAAGACGATGTCATCAACATGATCCCAAGACATTTGTTCAAAAGA 421  
 QY 432 TGCAATGATACACTACTGCTGACGCTCAACAAACCTTGCAATTTACATGATCCTCT 491  
 DB 422 TGCAATGATACACTACTGCTGACGCTCAACAAACCTTGCAATTTACATGATCCTCT 481  
 QY 492 CTGCTCTCTCAAGATGTTGTTATTTTGGCATATACCTGCTGCTGTTAAGAAC 551  
 DB 482 CTGCTCTCTCAAGATGTTGTTATTTTGGCATATACCTGCTGCTGTTAAGAAC 541  
 QY 552 GGCCTTCGCTGCAATGAGAG--GAAATCATTAACAAGCGGTGACAA--GAGGCGCATCT 608  
 DB 542 GGCCTTCGCTGCAATGAGAGAGAAATTCATTAACAAGCGGTGACAA--GAGGCGCATCT 601  
 QY 609 TTTCTCATGCGTATTTGCTCCTAG--AAGGCTTTCTGAGGATCTAGTTGGGCT-TTCT 665  
 DB 602 TTTCTCATGCGTATTTGCTCCTAGAGGGGCTTTCTGAGGATCTAGTTGGGCTATCCT 661  
 QY 666 TTTCGGTTTGGGCGCATTTCA-GTTTCATGTGTGTACTATTCATATTTGTATTAAC 724  
 DB 662 TACTGGGTAGGGCGCAATTCACGTCCCATGTGTGTCTATTCATATTTGTATTAAC 721  
 QY 725 GCTTTCAACCAAGTGGGCAACAGAGAACTCACTGTATTAAC 770  
 DB 722 GG--TTCAACCAATTTGGGCAACAGAAACTCATCTGTATTAAC 765

RESULT 5  
 BF678971 997 bp mRNA linear EST 21-DEC-2000  
 LOCUS 602153608F1 NIH\_MGC\_83 Homo sapiens cDNA clone IMAGE:4294873 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BF678971  
 VERSION BF678971.1 GI:11952866  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens



Subtraccion by Bento Soares and M. Fatima Bonaldo.

Query Match	60.2%	Score 618.6	DB 1	Length 636
Best Local Similarity	99.2%	Pred. No. 5.7e-158		
Matches 632	Conservative	0	Mismatches 4	Indels 1
				Gaps 1

OY	391	GATGTCATCAACAATGATCCCAACAAACAATGTTCCAAAAGATGCAAAAGATACACTAC	450
Db	636	GATGTCATCAACAATGATCCCAACAAACAATGTTCCAAAAGATGCAAAAGATACACTAC	577
OY	451	CTGCAGCTCAACAACAACCTCTGCATATTTACATGACCTCTCTCTGCTCTCAAGAGTGTG	510
Db	576	CTGCAGCTCAACAAC-CCTCTGCATATTTACATGACCTCTCTCTGCTCTCAAGAGTGTG	518
OY	511	GTCATTTTGGCATATCATCACTGCTGTGCTTAAGAAGCGCTTTCGCTGCACAAAGGA	570
Db	517	GTCATTTTGGCATATCACTGCTGTGCTTGAAGAAGCGCTTTCGCTGCACAAAGGA	458
OY	571	GAGAAATCATTAACAGACGGTGGACAGAGAGGCCATCTTTTCTCATCGGTTATTTGCC	630
Db	457	GAGAAATCATTAACAGACGGTGGACAGAGAGGCCATCTTTTCTCATCGGTTATTTGCC	398
OY	631	TAGAAAGCGTCTTTGAGAGATCTAGTTGGGCTTTCTTTCTGGGTTTGGGCACTTCAATTC	690
Db	397	TAGAAAGCGTCTTTGAGAGATCTAGTTGGGCTTTCTTTCTGGGTTTGGGCACTTCAATTC	338
OY	691	TCATGTGTACTATTTCTATCATTAATTTGATAACGGTTTCAAAACAAGTGGGACACAGA	750
Db	337	TCATGTGTACTATTTCTATCATTAATTTGATAACGGTTTCAAAACAAGTGGGACACAGA	278
OY	751	GAACCTCACTCTGTATAATAACAATGAGGAATAGCCACAGCGGATCTCCAGACCAACATCTCTC	810
Db	277	GAACCTCACTCTGTATAATAACAATGAGGAATAGCCACAGCGGATCTCCAGACCAACATCTCTC	218
OY	811	CATGTTTTCCACAGCTCTCTCCAGCCCAACCCAAATAGCGCTCTATAGTATGACATCTCT	870
Db	217	CATGTTTTCCACAGCTCTCTCCAGCCCAACCCAAATAGCGCTCTATAGTATGACATCTCT	158
OY	871	GCGGCTTCAAGCCTTGTCCTCTCTTAAGTCTTTTAATCAGATAAATCCTCGAAGACT	930
Db	157	GCGGCTTCAAGCCTTGTCCTCTCTTAAGTCTTTTAATCAGATAAATCCTCGAAGACT	98
OY	931	TTGATTTTACAGCCCTGAGGCAAGTCTTCTTTGCTAGTTGAATATATGAGTGATGTTTTTC	990
Db	97	TTGATTTTACAGCCCTGAGGCAAGTCTTCTTTGCTAGTTGAATATATGAGTGATGTTTTTC	38
OY	991	CGTAATTAAGCAAAATTAATTTAAAAAAAAGAAAAGTT 1027	
Db	37	CGTAATTAAGCAAAATTAATTTAAAAAAAAGAAAAGTT 1	

RESULT 7	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
BF674593	902 bp	mRNA	linear	EST 21-DRC-2001								
	602337811F1	NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4274283 5', mRNA sequence.	BF674593	BF674593.1	GI:11948488	EST.	Homo sapiens (human)					
							Homo sapiens					
							Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Carcharia; Homnidae; Homo.					
							1 (bases 1 to 902)					
							NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .					
							National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)					
							Contact: Robert Strausberg, Ph.D.					
							Email: <a href="mailto:gsabbs-remail.nih.gov">gsabbs-remail.nih.gov</a>					
							Tissue Procurement: CLONETECH Laboratories, Inc.					
							cDNA Library Preparation: CLONETECH Laboratories, Inc.					

CDNA library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: L1CM1091 row: e column: 04  
High quality sequence stop: 629.  
<http://www.illumina.com>

**FEATURES**  
**source**

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4274283"
/lab_host="DH10B (T1 phage-resistant)"
/clone_id="NH_MGC_83"
/note="Organ: prostate; Vector: pDNR-LIB (Clontech);
Site 1: SfiI (ggccgcctcgcc); Site 2: SfiI
(ggcattatcgcc); 5' and 3' adaptor were used in cloning
as follows: 5' adaptor sequence: 5'-CAGCGCATTAATGACC-3'
and 3' adaptor sequence:
5'-ATCTAGAGCGCGAGCGCGCGCATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.4
kb (range 0.5-4.0 kb). 14/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA)."

```

Query Match	59.8%	Score 614	DB 2	Length 902
Best Local Similarity	95.5%	Prod. No. 1,1e-156		
Matches 707	Conservative	0	Mismatches 25	Indels 8
				Gaps 7
Qy	12	GGGCAAAAATTCAGGATTTGGTCCCGGAACAAAGTTATTCATTACAGATTAACAAC	71	
Db	2	GGGCAAAAATTCAGGATTTGGTCCCGGAACAAAGTTATTCATTACAGATTAACAAC	61	
Qy	72	TGATGCAGATGTTTCCCAAGCCOACTATTTTCTTCTTCAATTGCTGAACAAAGCT	131	
Db	62	TGATGCAGATGTTTCCCAAGCCOACTATTTTCTTCTTCAATTGCTGAACAAAGCT	121	
Qy	132	CGAAGAGCTGGAACATACCTTTGTCTCTTGAGAAATTTTCCGTGATGTTATTAAGAT	191	
Db	122	CGAAGAGCTGGAACATACCTTTGTCTCTTGAGAAATTTTCCGTGATGTTATTAAGAT	181	
Qy	192	ACATTGGCAAGAAAAGAAAGACCAACGATTTCTGAGATCCCAAGAGGGAACCCATGAA	251	
Db	182	ACATTGGCAAGAAAAGAAAGACCAACGATTTCTGAGATCCCAAGAGGGAACCCATGAA	241	
Qy	252	GACTTACGACATATACATGAAATTTAGCTGGTTTAAAGGTGCGAGAAAAGTCACTGGACAA	311	
Db	242	GACTTACGACATATACATGAAATTTAGCTGGTTTAAAGGTGCGAGAAAAGTCACTGGACAA	301	
Qy	312	AGAACACAGATGTATCGTCAGACATGAGAAATTAATMAAACGAGTTGATCAAGAAATTTAT	371	
Db	302	AGAACACAGATGTATCGTCAGACATGAGAAATTAATMAAACGAGTTGATCAAGAAATTTAT	361	
Qy	372	CTTTCTCTCCATTAAGAAGGATGTCTACACATGATATCCCAAGACAA--TTGTTCCAAAG	430	
Db	362	CTTTCTCTCCATTAAGAAGGATGTCTACACATGATATCCCAAGACAAATTTGTTCCAAAG	421	
Qy	431	ATGCAAAATGATATACCTACTGCTGACGCTCAGAAACACCTCTGCATATTACATGTACTCC	490	
Db	422	ATGCAAAATGATATACCTACTGCTGACGCTCAGAAACACCTCTGCATATTACATGTACTCC	481	
Qy	491	TCCTGCTCTCAAGATGTGTTATTTTGGCATCATCACTGCTGTCTGCTTAGAAGAA	550	
Db	482	TCCTGCTCTCAAGATGTGTTATTTTGGCATCATCACTGCTGTCTGCTTAGAAGAA	541	
Qy	551	CGGCTTTCTGCTGCATGAGAGAAAT--CATTAACAGAGGTGACCAAGAGGCCATCTT	609	
Db	542	CGG--TTTCTGCTGCATGAGAGACATCATTAACAGAGGTGACCAAGAGG--CATCTT	599	
Qy	610	TTCTCATCGTTATTTGCTTCCAGAACGCTCTTCTGAGAGATC--TAGTTGGGCTTTCTTT	667	

Db 600 TTCCTCATCGTATAGTCCCTAGAAAGCGTCTTCTGAAGATCCAGTGGGGCTCTCTT 659

Qy 668 CTGGGT-TTGGGCGATTTCAGTTCTCATG-CTACTATCTATCTATTATGATTAACG 725

Db 660 CTGGGTCTGGGGCCATTCAGTTCTCATGTTGTTCATTCCTATCTATTATGATTAAC 719

Qy 726 GTTTTCAACCAAGTGGGCAC 745

Db 720 GGTTTAAACCAAGGCGCAC 739

RESULT 8  
AI685999/c  
LOCUS  
DEFINITION  
638 bp mRNA linear EST 27-MAY-1999  
tc91a04.x1 NCI CGAP Pr28 Homo sapiens CDNA clone IMAGE:2248878 3' similar to gb:M1321 T-CELL RECEPTOR GAMMA CHAIN C REGION (HUMAN) ; mRNA sequence.

ACCESSION  
AI685999  
EST.  
AI685999.1 GI:4897293

VERSION  
KEYWORDS  
Homo sapiens (human)

SOURCE  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
NCI CGAP htcp://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
Unpublished (1997)

JOURNAL  
COMMENT  
Contact: Robert Straubeberg, Ph.D.  
Email: cgaps-remail.nih.gov  
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
www-bio.lnl.gov/bdip/image/image.html  
Seq primer: -40UP from GIDCO  
High quality sequence stop: 453.

FEATURES  
SOURCE  
location/Qualifiers  
1..638  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2248878"  
/sex="male"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_id="NCI CGAP Pr28"  
/note="Organ: prostate; Vector: pT7AD-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI CGAP Pr22 was prepared, and 86 clones were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonids 985608-986759, 1101192-1101959, and 1217928-1220615). Subtraction by Bento Soares and M. Fatima Bonaudo."

ORIGIN  
Query Match 59.8%; Score 613.8; DB 1; Length 638;  
Best Local Similarity 98.6%; Pred. No. 1.2e-156;  
Matches 629; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

Qy 391 GATGTCATCAACGATCCCAAGACATTTTC-AAAAGATCAATGATTAACACTACT 449

Db 638 GATGTCATCAACGATCCCAAGACATTTTC-AAAAGATCAATGATTAACACTACT 579

Qy 450 GCTGAGCTCAACACCTGTGATATTATACATGACTCTCTGCTCCCAAGAGTGT 509

Db 578 GTCGAGCTCACAAACCTCTGATATTAACATGACTCATCTCTCTCAAGAGTGT 519

Qy 510 GGTCTATTTTGCATATACACCTGCTCTGCTTAAGAAAGCGTTTCGTCGAATGG 569

Db 518 GGTCTATTTTGCATATACACCTGCTCTGCTTAAGAAAGCGTTTCGTCGAATGG 459

Qy 570 AGAAGATCATTAACAGACGGTGGCACAAGAGGCGATCTTCTCATCGATTGTTC 629

Db 458 AGAAGATCATTAACAGACGGTGGCACAAGAGGCGATCTTCTCATCGATTGTTC 399

Qy 630 CTGAAGCGCTCTGAGGATCTAGTTGGGCTTCTTGGGTTTGGGCATTTCAGTT 669

Db 398 CTGAAGCGCTCTGAGGATCTAGTTGGGCTTCTTGGGTTTGGGCATTTCAGTT 339

Qy 690 CTGATGTGTACTATTTCTATCTATTATGATTAACGTTTCAACAGTGGGCACACAG 749

Db 338 CTGATGTGTACTATTTCTATCTATTATGATTAACGTTTCAACAGTGGGCACACAG 279

Qy 750 AGAAGCTCACTGTATATTAACATGAGAAATAGCCAGGCGATCTCCAGACCAATCTCT 809

Db 278 AGAAGCTCACTGTATATTAACATGAGAAATAGCCAGGCGATCTCCAGACCAATCTCT 219

Qy 810 CCATGTTTTCACAGCTCTCCAGCCCAACCAATAGCGCTGATATGATGACATCC 869

Db 218 CCATGTTTTCACAGCTCTCCAGCCCAACCAATAGCGCTGATATGATGACATCC 159

Qy 870 TCGGCTTCTAGCGCTTGTCCCTCTTCTTATGATTTTATGATTAACGCTCGAAGCC 929

Db 158 TCGGCTTCTAGCGCTTGTCCCTCTTCTTATGATTTTATGATTAACGCTCGAAGCC 99

Qy 930 TTTCAATTTTACAGCGCTGAGAGATCTTCTGATGATTTATGATGATTTT 989

Db 98 TTTCAATTTTACAGCGCTGAGAGATCTTCTGATGATTTATGATGATTTT 39

Qy 990 CCGTAAATTAAGCAAAATTAATTTTAAAAATGAAGAATT 1027

Db 38 CCGTAAATTAAGCAAAATTAATTTTAAAAATGAAGAATT 1

RESULT 9  
CN645427  
LOCUS  
DEFINITION  
ILLUMIGEN MCO 23989 Katze\_MSP Macaca mulatta cDNA clone  
IBIDW:10243 5' similar to Bases 1 to 948 highly similar to human  
TRG@ (Hs:385086), mRNA sequence.

ACCESSION  
CN645427  
EST.  
CN645427.1 GI:47158870

VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Macaca mulatta (rhesus monkey)

REFERENCE  
AUTHORS  
Magnees C.L., Fellin P.C., Thomas M.J., Korth M.J., Agy M.B., Prohl S.C., Fitzibbon M., Scherer C.A., Miner D.G., Katze M.G. and Iadonato S.P.  
Analysis of the Macaca mulatta transcriptome and the sequence divergence between Macaca and human  
Genome Biol. 6 (7), R60 (2005)  
15998449

JOURNAL  
PUBMED  
Contact: C. Magnees  
Illumigen Biosciences Inc.  
2203 Airport Way S, Suite 450, Seattle, WA 98134, USA  
Tel: 2063780400  
Fax: 2063780408  
Email: cmagnees@illumigen.com  
Sequenced on 2004.02.24. 794 Q20 bases.  
PCR Primers  
FORWARD: CCCTCACTAAAGGAACAAA  
BACKWARD: CACTTAGGCGCAATGGGTA  
Insert Length: 959 Std Error: 0.00

Plate: CU000136 row: G column: 03  
Seq primer: CCTCCTCACTAAAGGAAACAAA  
POLYA=yes.

## FEATURES

source

## Location/Qualifiers

1..959  
/organism="Macaca mulatta"  
/mol\_type="mRNA"  
/strain="Indian"  
/db\_xref="taxon:9544"  
/clone="IBIUM:10243"  
/sex="male"  
/cell\_type="mononuclear lymphocyte"  
/dev\_stage="adult"  
/lab\_host="E. coli SOLR"  
/clone\_1lb="Katze\_MMSB"  
/note="Organ: spleen; Vector: Uni-ZAP XR; Site: 1; Site 2: Xho I; Created from Strategene ZAP-CDNA Synthesis kit (catalog #200400) and ZAP-CDNA GigaPack III Gold Cloning Kit (Catalog #200450)"

## ORIGIN

Query Match 58.9%; Score 605.2; DB 7; Length 959;

Best Local Similarity 81.3%; Pred. No. 2.9e-154;  
Matches 758; Conservative 0; Mismatches 78; Indels 96; Gaps 1;

189 GATACATTGGCAAGAAAGAGACACGATTTCTGGATCCAGAGGGGAAACACCAT 248  
1 GATACATTGGCAAGAAAGAGACACGATTTCTGGATCCAGAGGGGAAACACCAT 60  
249 GAAGCTAAGCAATCATGAAGATTTAGTGTGTTACGGTGCAGAGAAAGTCACTGGA 308  
61 GAAGCTAAGCAATCATGAAGATTTAGTGTGTTACGGTGCAGAGAAAGTCACTGGA 120  
309 CAAGAAGACAGATGTATGTCAGACATGAGATATATATAAAGCGAGTTGATCAAGAAAT 368  
121 TAAAGACAGAGATGTATGTCAGACATGAGATATATATAAAGCGAGTTGATCAAGAAAT 180  
369 TATCTTTCTCCAAATAAGACGAGATG----- 395  
181 TATCTTTCTCCAAATAAGACAGATGTACACACAGTGTATCCAAAGACAGTTTTCAAA 240  
396 ----- 395  
241 AGACGCAATGATGTACACACAGTGTATCCAAAGCAATTAATTCAGAAATGACAGATGA 300  
396 ---CATCAATGATGTATCCAAAGCAATGTTTCAAAAGATGCAATGATCACTACTGCT 452  
301 TGACGCAAGTGTATCCCAAGACATTAATTCAGAAATGCAATGATGCACTACTGCT 360  
453 GCAGCTCAAAACACCTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 512  
361 GCAGCTCAAAACACCTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 420  
513 CTATTTTGGCATCAACCTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 572  
421 CTATTTTGGCATCAACCTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 480  
573 GAATCATTAAGACAGTGTGCAACAAGAGCCATCTTTTCTCATCGGTTATGTCCTTA 632  
481 GAGATTTGAACAGAGTGTGCAACAAGAGCCATCTTTTCTCATCGGTTATGTCCTTA 540  
633 GAAGGCTTTTGAAGATCTAGTTGGGCTTTCTTTCTGGGTTTGGGCAATTCAGTTTCTC 692  
541 GAAGGCTTTTGAAGATCTAGTTGGGCTTTCTTTCTGGGTTTGGGCAATTCAGTTTCTC 600  
693 ATGTGTGATCTATTTTATATATTTATTTATACGGTTTTCAAACGATGGGCAACAGAGA 752  
601 ATGTGTGATCTATTTTATATATTTATTTAAATGTTTTCAAACGATGGGCAACAGAGA 660  
753 ACCTCACTCTGATTAAGCAATGAGGATAGCAAGGCGCATCTCCAGACCAATCTCTCCA 812  
661 ACCTCACTCTGATTAAGCAATGAGGATAGCAAGGCGCATCTCCAGACCAATCTCTCCA 720

Qy 813 TGTTTTCCAGAGCTCTCTCCAGCCCAACCAATAGCGCTGTATAGTGTAGACATCTGCG 872  
Db 721 TAGTCTCCAGAGCTCTCTCCAGCCCAACCAATAGCGCTGTATAGTGTAGACATCTGCG 780  
Qy 873 GGGTTTACGCTGTGCTGCTCTTGTAGTGTCTTTATACAGATTAATGCTGGAAGCCTTT 932  
Db 781 AGCTTTACGCTGTGCTGCTCTTGTAGTGTCTTTATACAGATTAATGCTGGAAGCCTTT 840  
Qy 933 CATTTTACAGCGCTGGAAGAGCTTTTGTAGTGTATTAATGAGTGTGTGTTTCCG 992  
Db 841 CATTTTACAGTGTGGAAGAGCCTTTTGTAGTGTATTAATGAGTGTGTGTTTCCG 900  
Qy 993 TAATACCAAAATTAATTTAAAAATGAAA 1024  
Db 901 AATAGCCAAAATTAATTTAAAAATGAAA 932

## RESULT 10

LOCUS

BC217853 703 bp mRNA linear EST 21-APR-2001

DEFINITION RST37574 Atherys RAGE Library Homo sapiens cDNA, mRNA sequence.

ACCESSION BC217853  
VERSION BC217853.1 GI:13743874  
KEYWORDS EST.

SOURCE

ORGANISM Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 703)  
Harrington, J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R.,  
Cain, S., Leventhal, C., Thornton, M., Ramchandran, R.,  
Whittington, J., Lerner, L., Costanzo, D., McEligott, K., Booser, S.,  
Mays, R., Smith, B., Veloso, N., Kliska, A., Hess, J., Cochran, K., Lo, K.,  
Offenbacher, J., Danzig, J. and Ducar, M.  
Creation of genome-wide protein expression libraries using random  
activation of gene expression  
Nat. Biotechnol. 19 (5), 440-445 (2001)

JOURNAL

PUBMED  
11329013

COMMENT

Contact: Scott J. Cain  
Atherys, Inc.  
3201 Carnegie Ave, Cleveland, OH 44115, USA  
Tel: 216 431 9900  
Fax: 216 361 9596  
Email: scain@atherys.com

High quality sequence stop: 548.

FEATURES

Location/Qualifiers  
1..703  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_1lb="Atherys RAGE Library"  
/note="See 'Creation of Genome-wide Protein Expression'  
Libraries using Random Activation of Gene Expression', the  
Nature Biotechnology, in press. Note that even though the  
cell type indicated is HT1080, since a random activation  
method was used, these sequence tags are not necessarily  
expressed in HT1080 under normal circumstances."

## ORIGIN

Query Match 57.5%; Score 590.2; DB 2; Length 703;

Best Local Similarity 97.3%; Pred. No. 3.5e-150;  
Matches 621; Conservative 0; Mismatches 14; Indels 3; Gaps 2;

Qy 391 GATGTATCAATGATGCCAAAGACATTTGTTCAAAAGATGCAATGATGATCACTACTG 450  
Db 53 GATGTATCAATGATGCCAAAGAC-ATTGGTCAAAAGATGCAATGATGATCACTACTG 111  
Qy 451 CTGAGCTCAAAACACCTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 510  
Db 112 CTGAGCTCAAAACACCTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 171

QY 511 GTCTATTTTCCATCATCACCCTGCTGCTTCTTGAAGAACGGCTTCTCTGCAATGGA 570  
Db 172 GTCTATTTTCCATCATCACCCTGCTGCTTCTTGAAGAACGGCTTCTCTGCAATGGA 231  
QY 571 GAGAAATCATTAACAGACGGTGGCACAAGAGGCCATCTTTTCTCATCGTTATGTCCTC 630  
Db 232 GAGAAATCATTAACAGACGGTGGCACAAGAGGCCATCTTTTCTCATCGTTATGTCCTC 291  
QY 631 TAGAAGGCTTCTGAGGATCTAGTGGGCTTCTTCTGGGTTTGGGCAATTCAGTTC 690  
Db 292 TAGAAGGCTTCTAGAGGATCTAGTGGGCTTCTTCTGGGTTTGGGCAATTCAGTTC 351  
QY 691 TCATGTGTACTATTTCTATCATTTATTTGATTAACGTTTTCACACAGTGGCACAAGA 750  
Db 352 TCATGTGTACTATTTCTATCATTTATTTGATTAACGTTTTCACACAGTGGCACAAGA 411  
QY 751 GAACCTCACTCTGTAATTAACAGATGAGAAATGACAGGCGCATCTCCAGCAACAATCTTC 810  
Db 412 GAACCTCACTCTGTAATTAACAGATGAGAAATGACAGGCGCATCTCCAGCAACAATCTTC 471  
QY 811 CAGGTTTCCACAGCTCTCTCCAGCAACCCAAATAGGCGCTGCTATGTGATGACATCTC 870  
Db 472 CAGGTTTCCACAGCTCTCTCCAGCAACCCAAATAGGCGCTGCTATGTGATGACAGCTC 531  
QY 871 GCGGCTCTAGCCTTGTCCCTCTCTTATGTTTCTTATCATGATGATGCTGGAAGCCT 930  
Db 532 GCGGCTCTAGCCTTGTCCCTCTCTTATGTTTCTTATCATGATGATGCTGGAAGCCT 591  
QY 931 TTCAATTTTACAGCCCTGGAAGAGCTCTTCTTGTAGTTGATTTATGTTG-1GTGTTT 988  
Db 592 TTCAATTTTACAGCCCTGGAAGAGCTCTTCTTGTAGTTGATTTATGTTGTTT 651  
QY 989 TCCGTAATAGCAAAATTAATTTAAAAAATGAAAAAT 1026  
Db 652 CCCGTAATAGCAAAATTAATTTAAAAAATGAAAAAT 689

RESULT 11  
A1768834/c 629 bp mRNA linear EST 21-DEC-1999  
LOCUS A1768834  
DEFINITION w03h04.x1 NCI CGAP K1d12 Homo sapiens CDNA clone IMAGE:2401783 3,  
similar to gb:M33231 T-CELL RECEPTOR GAMMA CHAIN C REGION (HUMAN);  
mRNA sequence.

ACCESSION A1768834.1 GI:5235343  
VERSION A1768834  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 629)  
AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index

JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaps-remail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/BLM at:  
www-bio.llnl.gov/bbrp/image/image.html  
Insert length: 1658 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 428.  
Location/Qualifiers  
1. .629  
/organism="Homo sapiens"  
/mol\_type="mRNA"

/db\_xref="taxon:9606"  
/clone="IMAGE:2401783"  
/issue\_type="2 pooled tumors (clear cell type)"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP K1d12"  
/note="Organ: Kidney; Vector: pT7T3D-Pac (Pharmacia) with  
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;  
Plasmid DNA from the normalized library NCI CGAP K1d5 was  
prepared, and as circles were made in vitro. Following HAP  
purification, this DNA was used as tracer in a subtractive  
hybridization reaction. The driver was PCR-amplified cDNAs  
from a pool of 5,000 clones made from the same library  
(cloneids 1323912-1325831, 1471368-1472903 and  
1492104-1493255). Subtraction by Bento Soares and M.  
Fatima Bonaldo."

ORIGIN  
Query Match 57.4%; Score 589.6; DB 1; Length 629;  
Best Local Similarity 97.9%; Pred. No. 5e-150;  
Matches 606; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

QY 406 GATCCCAAGACAAATGTTTCAAAAAGATGCAATGATACACTACTGCTGCACTCAACAAC 465  
Db 621 GATCCCAAGACAAATG-TCATATGATGCAAAATGATACANTCCTGATGAGCTCACAAAC 563  
QY 466 ACCTGCAATTAATACATGATACCTCTCTGCTCTCTCAAGATGATGCTATTTTGGCATC 525  
Db 562 ACTNCTGCAATTAATACATGATACCTCTCTCTCTCTCAAGATGATGCTATTTTGGCATC 503  
QY 526 ATCACTGCTGCTGCTTGAAGAGAGCGCTTCTGCTGCAATGAGAGAAATCATACAG 585  
Db 502 ATCACTGCTGCTGCTTGAAGAGAGCGCTTCTGCTGCAATGAGAGAAATCATACAG 443  
QY 586 ACCGTGGACAGAGAGCCCATCTTTTCTCATGCTTATGTTTCCCTTGAAGCGCTTCTG 645  
Db 442 ACCGTGGACAGAGAGCCCATCTTTTCTCATGCTTATGTTTCCCTTGAAGCGCTTCTG 383  
QY 646 AGGATCAGTTGGGCTTTCTTCTGGGTTTGGGCAATTCAGTTTCATGTTGATCTAT 705  
Db 382 AGGATCAGTTGGGCTTTCTTCTGGGTTTGGGCAATTCAGTTTCATGTTGATCTAT 323  
QY 706 TCTATCATTAATTTATTAACCGTTTTCACACAGTGGGACACAGAACTCACTGTGA 765  
Db 322 TCTATCATTAATTTATTAATAGTTTTCACACAGTGGGACACAGAACTCACTGTGA 263  
QY 766 ATAAATGAGAGAAATGACCGGCGATCTCAGACCAATCTCTCATGTTTTCACAGC 825  
Db 262 ATAAATGAGAGAAATGACCGGCGATCTCAGACCAATCTCTCATGTTTTCACAGC 203  
QY 826 TCCGCCAGCAACCCAAATAGGCGCTGCTATGATGATGATCCTGCGGCTTACGCTT 885  
Db 202 TCCGCCAGCAACCCAAATAGGCGCTGCTATGATGATGATCCTGCGGCTTACGCTT 143  
QY 886 GTCCCTCTCTTATGTTCTTAAATCAGATTAAGTCTGGAAGCGCTTCAATTTACAGCC 945  
Db 142 GTCCCTCTCTTATGTTCTTAAATCAGATTAAGTCTGGAAGCGCTTCAATTTACAGCC 83  
QY 946 CTGAAGCAGTCTTCTTCTGCTAGTTGAATTAATGTTGTTGTTTTCGTAATTAACAAAT 1005  
Db 82 CTGAAGCAGTCTTCTTCTGCTAGTTGAATTAATGTTGTTGTTTTCGTAATTAACAAAT 23  
QY 1006 AAATTTAAAAAATGAAAA 1024  
Db 22 AAATTTAAAAAATGAAAA 4

RESULT 12  
CA449324/c 700 bp mRNA linear EST 08-NOV-2002  
LOCUS CA449324  
DEFINITION UT-H-E11-ayc-k-17-0-UT.s1 NCI CGAP E11 Homo sapiens CDNA clone  
CA449324  
ACCESSION CA449324  
VERSION CA449324.1 GI:24813744

KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 700)  
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.  
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
TITLE Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Dr. Jose Mercuende  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be obtained  
from Dr. M. Bento Soares, bento-soares@uiowa.edu  
The following repetitive elements were found in this cDNA  
sequence: 1-29, >AT richLow\_complexity  
Seq primer: M13 FORWARD  
POLYA=yes

FEATURES  
Source  
1..700  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-H-B11-ayc-k-17-0-UI"  
/tissue\_type="Chondrosarcoma"  
/dev\_stage="Adult"  
/lab\_host="DH10B (Life Technologies)"  
/note="Organ: Left Pelvis; Vector: pRTT3-Pac (Pharmacia)  
with a modified polylinker; Site 1: EcoR I; Site 2: Not I;  
NCI CGAP\_B11 is a normalized cDNA library containing the  
following tissue(s): Chondrosarcoma. The library was  
constructed according to Bonaldo, Lennon and Soares,  
Genome Research, 6:791-806, 1996. First strand cDNA  
synthesis was primed with an oligo-dT primer containing a  
Not I site. Double stranded cDNA was ligated to an EcoR I  
adaptor, digested with Not I, and cloned directionally  
into pRTT3-Pac vector. The oligonucleotide used to prime  
the synthesis of first-strand cDNA contains a library tag  
sequence that is located between the Not I site and the  
(dT)18 tail. The sequence tag for this library is  
AACTGTGCAC  
TAG\_TISSUE=chondrosarcoma  
TAG\_LIB=UI-H-B11  
TAG\_SEQ=ACACTTGCAC"

ORIGIN  
Query Match 56.9%; Score 584; DB 6; Length 700;  
Best Local Similarity 95.9%; Pred. No. 1.7e-148;  
Matches 629; Conservative 0; Mismatches 5; Indels 22; Gaps 2;  
391 GATGTCATCACAATGATCCCAAGACATTTGCAAAAGATGCAAT----- 438  
661 GATGTCATCACAATGATCCCAAGACATTTGCAAAAGATGCAATGTCGAA 602  
439 -----GATACACTACTGCTGCAGCTCACAACACCTTGCAATTTAATGTAAGCTC 489  
601 TTTCTCTTAATACACTACTGCTGCAGCTCACAACACCTTGCAATTTAATGTAAGCTC 542  
490 CTCCTGCTCTTC-AGAGTGTGTCTATTTTGGCATCTACCTGCTGCTTGAAG 548  
541 CTCCTGCTCTTCNAAGAGTGTGTCTATTTTGGCATCTACCTGCTGCTTGAAG 482  
549 AACGGCTTTCTGCTCAATGAGAGAAATCATATACAGACGGTGCACAAGAGGCATCT 608  
481 AACGGCTTTCTGCTCAATGAGAGAAATCATATACAGACGGTGCACAAGAGGCATCT 422  
609 TTTCTCATCGGTATATGTCTCCTAGAGCGTCTTCTGAGATCTAGTTGGGCTTTCTTTC 668

Db 421 TTTCTCATCGGTATATGTCTCCTAGAGCGTCTTCTGAGATCTAGTTGGGCTTTCTTTC 362  
Qy 669 TGGATTGTGGGCATTTTCAGTTCTCATGTGTACTATTCATATTAATTAACGGTT 728  
Db 361 TGGATTGTGGGCATTTTCAGTTCTCATGTGTACTATTCATATTAATTAATGTT 302  
Qy 729 TTCAACACGATGGGCACACAGAGAACTCACTGTATATACATGAGAAATGCCACGG 788  
Db 301 TTCAACACGATGGGCACACAGAGAACTCACTGTATATACATGAGAAATGCCACGG 242  
Qy 789 CGATCTCCAGACCAATCTCCTCATATGTTTCCACAGCTCTCCAGCCCAACCAATAGCG 848  
Db 241 CGATCTCCAGACCAATCTCCTCATATGTTTCCACAGCTCTCCAGCCCAACCAATAGCG 182  
Qy 849 CCGCTATATGATAGACATCTCGGCTTCTAGCTTTCCTCTTCTAGTGTCTTTAA 908  
Db 181 CCGCTATATGATAGACATCTCGGCTTCTAGCTTTCCTCTCTTATGTTCTTTAA 122  
Qy 909 TCAATTAATGCTGTGCTTTTCCGTATATAGCAAAATTTTAAATTAATGAAAA 1024  
Db 121 TCAATTAATGCTGTGCTTTTCCGTATATAGCAAAATTTTAAATTAATGAAAA 62  
Qy 969 TGAATTAATGCTGTGCTTTTCCGTATATAGCAAAATTTTAAATTAATGAAAA 1024  
Db 61 TGAATTAATGCTGTGCTTTTCCGTATATAGCAAAATTTTAAATTAATGAAAA 6

RESULT 13  
BE326754/c  
LOCUS BE326754.1 582 bp mRNA linear EST 14-JUN-2000  
DEFINITION h64d05.x1 NCI CGAP\_Kid11 Homo sapiens cDNA clone IMAGE:313257.3  
similar to gb:M1321 T-CELL RECEPTOR GAMMA CHAIN C REGION (HUMAN);,  
mRNA sequence.  
ACCESSION BE326754  
VERSION BE326754.1 GI:9200530  
KEYWORDS EST.  
ORGANISM Homo sapiens (human)  
SOURCE Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 582)  
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.  
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
TITLE Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
DNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CCAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL, send email to:  
info@image.jhl.gov  
Seq primer: -40UP from Gldco  
High quality sequence stop: 447.  
Location/Qualifiers  
1..582  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:313257"  
/lab\_host="DH10B"  
/note="Organ: Kidney; Vector: pRTT3-Pac (Pharmacia) with  
a modified polylinker; Site 1: Not I; Site 2: Eco RI;  
Plasmid DNA from the normalized library NCI CGAP\_Kid11 was  
prepared, and ss circles were made in vitro. Following HAP  
purification, this DNA was used as tracer in a subcloned  
hybridization reaction. The driver was PCR-amplified cDNAs

## ORIGIN

from a pool of 5,000 clones made from the same library  
(cloneId: 1322376-1323911, 1456007-1456775, and  
1500552-1502855). Subtraction by Bento Soares and M.  
Fatima Bonaldo.

Query Match 56.5%; Score 580.4; DB 2; Length 582;  
Best Local Similarity 99.8%; Pred. No. 1.6e-147;  
Matches 581; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 439 GATACACTACTGCTGCGCTCACAACAACCTTGCATATTAACATGTAACCTCTCTGCTC 498  
Db 582 GATACACTACTGCTGCGCTCACAACAACCTTGCATATTAACATGTAACCTCTCTGCTC 523  
Qy 499 CTGAAGGTGTGCTATTTTGGCATCATACCGTGTGCTGCTTGAAGAAGGCTTTC 558  
Db 522 CTGAAGGTGTGCTATTTTGGCATCATACCGTGTGCTGCTTGAAGAAGGCTTTC 463  
Qy 559 TGCTGCATGGAAGAAATCATACAGACGGTGGCACAAGAGGCACTCTTTCCTCATC 618  
Db 462 TGCTGCATGGAAGAAATCATACAGACGGTGGCACAAGAGGCACTCTTTCCTCATC 403  
Qy 619 GGTATTTGCTCCAGAAAGCGTCTTCTGAGATCTAGTTGGGCTTCTTCTGGTTGGG 678  
Db 402 GGTATTTGCTCCAGAAAGCGTCTTCTGAGATCTAGTTGGGCTTCTTCTGGTTGGG 343  
Qy 679 CCATTTCAGTCTCATGCTGCTACTATTCATCATTTTGTATTAAGCGTTTTCACAACAG 738  
Db 342 CCATTTCAGTCTCATGCTGCTACTATTCATCATTTTGTATTAAGCGTTTTCACAACAG 283  
Qy 739 TGGGCAACAGAGAACTCACTCTGTATTAACAATGAGAAATGACCGGATCTCAG 798  
Db 282 TGGGCAACAGAGAACTCACTCTGTATTAACAATGAGAAATGACCGGATCTCAG 223  
Qy 799 CACCAATCTCTCATGTTTTCACAGCTCTCCAGCAACCCAAATAAGCGCTGCTATAG 858  
Db 222 CACCAATCTCTCATGTTTTCACAGCTCTCCAGCAACCCAAATAAGCGCTGCTATAG 163  
Qy 859 TGTAGACATCTGGGCTTCTAGGCTTGTCCCTCTTATGTTCTTAAATCAATTAAT 918  
Db 162 TGTAGACATCTGGGCTTCTAGGCTTGTCCCTCTTATGTTCTTAAATCAATTAAT 103  
Qy 919 GCTCGAAGCTTTCATTTTACAGCGCCCTGAAGAGCTTCTTGTAGTTGAATTAAT 978  
Db 102 GCTCGAAGCTTTCATTTTACAGCGCCCTGAAGAGCTTCTTGTAGTTGAATTAAT 43  
Qy 979 GGTGTGTTTTCCTTAATAAGCAAAATTAATTTAAAAAATG 1020  
Db 42 GGTGTGTTTTCCTTAATAAGCAAAATTAATTTAAAAAATG 1

RESULT 14 613 bp mRNA linear EST 16-OCT-2001  
BI905966 60306285F1 NIH\_MGC\_118 Homo sapiens cDNA clone IMAGE:5212261 5',  
LOCUS mRNA sequence.  
ACCESSION BI905966.1 GI:16168605  
VERSION BI905966  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
TITLE Homidae; Homo.  
JOURNAL NIH-MGC <http://mgc.nci.nih.gov/>.  
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)

DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
<http://image.lnl.gov>  
Plate: LLM11532 row: 0 column: 14  
High quality sequence stop: 578.  
Location/Qualifiers  
1. 613  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5212261"  
/feature\_type="Leukocyte"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_118"  
/note="Vector: pCMV-Sport6; Site\_1: NotI; Site\_2: EcoRV  
(destroyed); RNA source Leukocytes from anonymous pool of  
non-activated adult donors. Library is oligo-dT primed  
and directionally cloned (EcoRV site is destroyed upon  
cloning). Average insert size 1.7 kb, insert size range  
1.2-3.3 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 027. Note:  
this is a NIH\_MGC Library."

FEATURES  
source

## ORIGIN

Query Match 56.4%; Score 578.8; DB 3; Length 613;  
Best Local Similarity 99.7%; Pred. No. 4.5e-147;  
Matches 580; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 443 CACTACTGCTGACGCTCACAACAACCTTGCATATTAACATGTAACCTCTCTGCTCA 502  
Db 1 CACTACTGCTGACGCTCACAACAACCTTGCATATTAACATGTAACCTCTCTGCTCA 60  
Qy 503 AGAGTGTGCTATTTTGGCATCATCACTGCTGTCTGCTTAAAGAAGCGCTTTCGCT 562  
Db 61 AGAGTGTGCTATTTTGGCATCATCACTGCTGTCTGCTTAAAGAAGCGCTTTCGCT 120  
Qy 563 GCAATGAGAGAAATCATTAACAAGGTGGCAACAAGAGGCACTTTTCTCATGCGGT 622  
Db 121 GCAATGAGAGAAATCATTAACAAGGTGGCAACAAGAGGCACTTTTCTCATGCGGT 180  
Qy 623 ATTGTCCTTGAAGCGTCTTCTGAGATCTAGTTGGGCTTCTTCTTGCGGTTGGGCAT 682  
Db 181 ATTGTCCTTGAAGCGTCTTCTGAGATCTAGTTGGGCTTCTTCTTGCGGTTGGGCAT 240  
Qy 683 TTGAGTTCATGTTGTACTATTTATCATTAATTGTATACGATTTTCAACCAAGTGG 742  
Db 241 TTGAGTTCATGTTGTACTATTTATCATTAATTGTATACGATTTTCAACCAAGTGG 300  
Qy 743 CACACAGAGAACTCACTCTGTATTAACAATGAGAAATGACCAAGGATCTCCAGACC 802  
Db 301 CACACAGAGAACTCACTCTGTATTAACAATGAGAAATGACCAAGGATCTCCAGACC 360  
Qy 803 AATCTCTCATGTTTTCACAGCTCTCCAGCAACCAATGAGCGCTCTATAGTGA 862  
Db 361 AATCTCTCATGTTTTCACAGCTCTCCAGCAACCAATGAGCGCTCTATAGTGA 420  
Qy 863 GACATCTGGGCTTCTAGGCTTGTCCCTCTCTTATGTTCTTAAATCAGATAATGCTCT 922  
Db 421 GACATCTGGGCTTCTAGGCTTGTCCCTCTCTTATGTTCTTAAATCAGATAATGCTCT 480  
Qy 923 GGAAGCTTTCATTTTACAGCGCTGAGAGAGCTTCTTGTGTAATTAATAGTGGTG 982  
Db 481 GGAAGCTTTCATTTTACAGCGCTGAGAGAGCTTCTTGTGTAATTAATAGTGGTG 540  
Qy 983 TGTATTTCCGTAAATAGCAAAATTAATTTAAAAAATGAATA 1024  
Db 541 TGTATTTCCGTAAATAGCAAAATTAATTTAAAAAATGAATA 582

RESULT 15  
BF681238

LOCUS BF681238 719 bp mRNA linear EST 21-DEC-2000  
 DEFINITION 602155502F1 NIH\_MGC\_83 Homo sapiens cDNA clone IMAGE:4296504 5',  
 mRNA sequence.  
 ACCESSION BF681238  
 VERSION BF681238.1 GI:11955133  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.  
 1 (bases 1 to 719)  
 NIH-MGC <http://mgs.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 CONTACT: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: CLONETECH Laboratories, Inc.  
 cDNA Library Preparation: CLONETECH Laboratories, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
 Plate: LHCMI149 row: c column: 01  
 High quality sequence stop: 673.  
 Location/Qualifiers  
 1..719  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4296504"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NIH MGC 83"  
 /note="Organ: prostate; Vector: pDNR-LIB (Clontech);  
 Site\_1: SfiI (ggcgctcgcc); Site\_2: SfiI  
 (ggcctatggcc); 5' and 3' adaptors were used in cloning  
 as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3',  
 and 3' adaptor sequence:  
 5'-ATTCTAGAGCGCGCGCCGACATG-dT(30)BN-3' (where B = A,  
 C, or G and N = A, C, G, or T). Average insert size 1.4  
 kb (range 0.5-4.0 kb). 14/15 colonies contained inserts  
 by PCR. This library was enriched for full-length clones  
 and was constructed by Clontech Laboratories (Palo Alto,  
 CA)."

## ORIGIN

Query Match 55.5%; Score 569.8; DB 2; Length 719;  
 Best Local Similarity 94.6%; Pred. No. 1.4e-144;  
 Matches 678; Conservative 0; Mismatches 27; Indels 12; Gaps 8;

QY 11 TGGGCAAAAAATCAAGGATTTGGTCCCGAACAAGCTTATCATTAAGATAAACAAC 70  
 DB 5 TGGGCAAAAAATCAAGGATTTGGTCCCGAACAAGCTTATCATTAAGATAAACAAC 64  
 QY 71 TTGATGAGATGTTTCCCAAGCCCACTATTTTCTTCAATTGCTGAAACAAGC 130  
 DB 65 TTGATGAGATGTTTCCCAAGCCCACTATTTTCTTCAATTGCTGAAACAAGC 124  
 QY 131 TCCAGAGGCTGGAAATACCTTTTCTTCTTGAATAATTTTCCCTGATGTTATTAAGA 190  
 DB 125 TCCAGAGGCTGGAAATACCTTTTCTTCTTGAATAATTTTCCCTGATGTTATTAAGA 183  
 QY 191 TACATTTGGCAAGAAAGAGAGCAACAGATTCTGGATCCCAAGAGGGGAAACACCATGA 250  
 DB 184 TACATTTGGCAAGAAAGAGAGCAACAGATTCTGGATCCCAAGAGGGGAAACACCATGA 243  
 QY 251 AGACTTAAGACATATCAATGAATTTAGCTGTTAACGGTGCAGAAAAAGTCACCTGACA 310  
 DB 244 AGACTTAAGACATATCAATGAATTTAGCTGTTAACGGTGCAGAAAAAGTCACCTGACA 303  
 QY 311 AAGAACACGATGTTCTGTCAGATGAGATTAATAAAGGAGTTGATCAAGAAATTA 370

DB 304 AAGAACACGATGTTCTGTCAGATGAGATTAATAAAGGAGTTGATCAAGAAATTA 363  
 QY 371 TCTTCTCCCAATTAAGACGAGATGTCACCAATGATCCCAAGCAATTTTCAAAAG 430  
 DB 364 TCTTCTCCCAATTAAGACGAGATGTCACCAATGATCCCAAGCAATTTTCAAAAG 423  
 QY 431 ATGCAATGATACCTACTGCTGAGCTCACAAACACCTTGTGATTTACATGTAACCTCC 490  
 DB 424 ATGCAATGATACCTACTGCTGAGCTCACAAACACCTTGTGATTTACATGTAACCTCC 483  
 QY 491 TCTGCTCTCAAGAGTGTGCTTA--TTTGGCATCATCACTGCTGCTT--AGAG 548  
 DB 484 TCTGCTCTCAAGAGTGTGCTTA--TTTGGCATCATCACTGCTGCTT--AGAG 543  
 QY 549 AACGGCTTTCTG--CTGCAATGAGAGAAATCATTAACAGAGGTGGAC--AAGAGGCCA 605  
 DB 544 AACGGCTTTCTG--CTGCAATGAGAGAAATCATTAACAGAGGTGGAC--AAGAGGCCA 603  
 QY 606 TCTTTTCTCATCGGTTA--TTGTCCTAGAAAGGCTCTTC--TGAGATCTAGTTGGGCTT 662  
 DB 604 TCTTTTCTCATCGGTTA--TTGTCCTAGAAAGGCTCTTC--TGAGATCTAGTTGGGCTT 663  
 QY 663 TCTTTCTGAGTTGGGCAATTTCAAGTCTCATGCTGCTATCTATCATTTATGT 719  
 DB 664 ACTTCTGGGTTGGCCATTTCAAGTCTTATGCTG--ACTATTTATCATTTATGT 717

Search completed: December 10, 2005, 23:23:07  
 Job time : 4086 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2005, 17:49:47 ; Search time 664 Seconds  
(without alignments)  
10308.192 Million cell updates/sec

Title: US-10-031-158B-13

Perfect score: 1027  
Sequence: 1 gggcagagtggtggcacaataa.....atttaaaaaatgcaaaagct 1027

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 333246308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: Genesegq1980s:\*  
2: Genesegq1990s:\*  
3: Genesegq2000s:\*  
4: Genesegq2001as:\*  
5: Genesegq2001bs:\*  
6: Genesegq2002as:\*  
7: Genesegq2002bs:\*  
8: Genesegq2003as:\*  
9: Genesegq2003bs:\*  
10: Genesegq2003cs:\*  
11: Genesegq2003ds:\*  
12: Genesegq2004as:\*  
13: Genesegq2004bs:\*  
14: Genesegq2005s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1027	100.0	1027	4	AAFS6401
2	1027	100.0	1027	4	ADBP75580
3	1027	100.0	1027	13	ADBP54862
4	1027	100.0	1027	14	ADW44400
5	1013.8	98.7	1395	12	ADL06443
6	984	95.8	1530	14	ADY16296
7	968.2	94.3	1799	6	ABZ35381
8	945.6	92.1	1418	6	AAU38828
9	915	89.1	2658	10	ADBP08798
10	906	88.2	1560	10	ADBP06987
11	905.6	88.2	1560	10	ADBP06987
12	902.2	87.8	1421	6	ADBP06987
13	902.2	87.8	1421	13	ADBP06987
14	902.2	87.8	1421	13	ADBP06987
15	838	81.6	1586	13	ADBP06987
16	838	81.6	1586	13	ADBP06987
17	838	81.6	1586	13	ADBP06987
18	838	81.6	1586	13	ADBP06987
19	835.2	81.3	2658	10	ADBP08798

20	821.6	80.0	1046	12	ADBP10449
21	795.8	77.5	1162	5	ABV27727
22	795.8	77.5	1162	5	ABV24153
23	795.8	77.5	1162	5	ABV24540
24	795.8	77.5	1162	5	ABV25162
25	795.8	77.5	1162	5	ABV21526
26	795.8	77.5	1162	5	ABV27345
27	795.8	77.5	1162	5	ABV24544
28	795.8	77.5	1162	5	ABV24851
29	660.2	64.3	865	14	ACU55249
30	619	60.3	924	5	ABV25092
31	601	58.5	786	13	ADBP55163
32	589	57.4	124933	14	ABV61187
33	582.6	56.7	765	14	ADY78562
34	522.8	50.9	1080	14	AAO37617
35	522.8	50.9	1080	14	ADY20108
36	516.4	50.3	1080	1	AAV21698
37	498.8	48.6	1316	5	ABV26485
38	498.8	48.6	1316	5	ABV25093
39	492	47.9	531	13	ADU12452
40	473.2	46.1	486	5	ABV25013
41	449.2	43.7	825	2	AAQ66888
42	449.2	43.7	825	9	ACF35988
43	446.8	43.5	784	5	ABV21896
44	444	43.2	511	2	AAV58549
45	444	43.2	511	2	AAV61291

#### ALIGNMENTS

RESULT 1  
AAFS6401  
ID AAF56401 standard; cDNA; 1027 BP.  
XX AAF56401;  
AC  
XX  
DT 12-APR-2001 (first entry)  
XX  
DE Human TCRgamma alternate reading frame protein coding sequence.  
XX  
KM Human; TARP; prostate cancer; breast cancer; immunotherapy;  
KM T cell receptor gamma alternate reading frame protein; TCRgamma; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200104309-A1.  
PD 18-JAN-2001.  
XX  
PF 12-JUL-2000; 2000MC-US019039.  
XX  
PR 13-JUL-1999; 99US-0143560P.  
PR 01-OCT-1999; 99US-0157471P.  
XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
PI Pastan I, Essand M, Lee B, Vasmatazis G, Wolfgang C;  
XX  
DR WPI: 2001-061050/09.  
XX  
PT P-PSDB; AAB66399, AAB66400.  
XX  
PS  
XX  
XX The present invention provides the protein and coding sequences of the  
XX human T cell receptor alternate reading frame protein (TARP). This  
XX protein is expressed in prostate and breast cancer cells at higher levels  
XX than normal and so can be used in the immunotherapy of these cancers, as  
XX well as their detection and prevention

```

XX Sequence 1027 BP; 307 A; 232 C; 200 G; 288 T; 0 U; 0 Other;
SQ Query Match 100.0%; Score 1027; DB 4; Length 1027;
Best Local Similarity 100.0%; Pred. No. 4,3e-290;
Matches 1027; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGCAAGAGTTGGGCAAAAAAATCAAGGTATTTGGTCCCGGAACAAGCTTATCATTAACA 60
DB 1 GGGCAAGAGTTGGGCAAAAAAATCAAGGTATTTGGTCCCGGAACAAGCTTATCATTAACA 60
QY 61 GATTAACAACCTTGTATGACAGATGTTTCCCGCAAGCCCATATTTTCTTCTCAATTGCT 120
DB 61 GATTAACAACCTTGTATGACAGATGTTTCCCGCAAGCCCATATTTTCTTCTCAATTGCT 120
QY 121 GAAACAAGCTCCAGAAAGCTGGAACATACCTTTGCTTCTTGAGAAATTTTCCCTGAT 180
DB 121 GAAACAAGCTCCAGAAAGCTGGAACATACCTTTGCTTCTTGAGAAATTTTCCCTGAT 180
QY 181 GTTATTAAATGATACATTGGCAAGAAAGAGCAACAGATTCTGGGATCCGAGAGGGG 240
DB 181 GTTATTAAATGATACATTGGCAAGAAAGAGCAACAGATTCTGGGATCCGAGAGGGG 240
QY 241 AACACCATGAAGACTTAACGACATACATGAAATTTAGCTGTTAACGGTCCAGAAAG 300
DB 241 AACACCATGAAGACTTAACGACATACATGAAATTTAGCTGTTAACGGTCCAGAAAG 300
QY 301 TCACCTGCAAGAAAGACAGATGTATCGTCAGACATGAGAAATTAATTAACGAGATTGAT 360
DB 301 TCACCTGCAAGAAAGACAGATGTATCGTCAGACATGAGAAATTAATTAACGAGATTGAT 360
QY 361 CAAGAATTAATCTTCTCTCAATAAAGAAGATGTCATCAGATGATCCCAAGACAT 420
DB 361 CAAGAATTAATCTTCTCTCAATAAAGAAGATGTCATCAGATGATCCCAAGACAT 420
QY 421 TGTTCAAAAAGATGCAATGATACACTACTGCTGCAACAACCTCTGCATATTAC 480
DB 421 TGTTCAAAAAGATGCAATGATACACTACTGCTGCAACAACCTCTGCATATTAC 480
QY 481 ATGTACCTCTCTGCTCTCTCAAGAGTGGTCTATTTTGGCATCATACCGCTGCTG 540
DB 481 ATGTACCTCTCTGCTCTCTCAAGAGTGGTCTATTTTGGCATCATACCGCTGCTG 540
QY 541 CTTAGAAAGACGGCTTTCTGCTGCAATGAGAAATCATTAACAGCGTGGCAACAAGA 600
DB 541 CTTAGAAAGACGGCTTTCTGCTGCAATGAGAAATCATTAACAGCGTGGCAACAAGA 600
QY 601 GGGCATCTTTTCTCTATGCGTTATTGTCCTTGAAGCGCTTTCTGAGGATCTAGTTGGGC 660
DB 601 GGGCATCTTTTCTCTATGCGTTATTGTCCTTGAAGCGCTTTCTGAGGATCTAGTTGGGC 660
QY 661 TTTCTTCTGGGTTGGGCAATTTAGTTCTCATGTTGTAATTTCTATCATTTATTTGTA 720
DB 661 TTTCTTCTGGGTTGGGCAATTTAGTTCTCATGTTGTAATTTCTATCATTTATTTGTA 720
QY 721 TAAACGTTTTCACCAAGTGGGCAACAGAAACCTCACTCTGTATTAACAATGAGAT 780
DB 721 TAAACGTTTTCACCAAGTGGGCAACAGAAACCTCACTCTGTATTAACAATGAGAT 780
QY 781 AGCCAGCGGATCTCCAGACCAATCTCTCCATGTTTCCACAGGCTCTCCAGCCAAACC 840
DB 781 AGCCAGCGGATCTCCAGACCAATCTCTCCATGTTTCCACAGGCTCTCCAGCCAAACC 840
QY 841 AAATAGCGGCTGTATGTTAGTACATCTCGGCTCTAGGCTTGCCCTCTTAAGTG 900
DB 841 AAATAGCGGCTGTATGTTAGTACATCTCGGCTCTAGGCTTGCCCTCTTAAGTG 900
QY 901 TTCTTTAATACAGATACTGCTGGAAGCTTTCATTTTAAAGCCCTGAAGAGCTTCT 960
DB 901 TTCTTTAATACAGATACTGCTGGAAGCTTTCATTTTAAAGCCCTGAAGAGCTTCT 960
QY 961 TTGCTAGTTGAATTAATGTTGTTGTTTCCGTAATTAAGCAAAATTAATTAATAATG 1020
DB 961 TTGCTAGTTGAATTAATGTTGTTGTTTCCGTAATTAAGCAAAATTAATTAATAATG 1020

```

```

DB 961 TTGCTAGTTGAATTAATGTTGTTGTTTCCGTAATTAAGCAAAATTAATTAATAATG 1020
QY 1021 AAAAGTT 1027
DB 1021 AAAAGTT 1027

RESULT 2
ADB75580
ID ADB75580 standard; cDNA; 1027 BP.
XX
AC ADB75580;
XX
XX 04-DEC-2003 (first entry)
XX
XX Prostate cancer marker cDNA.
XX
XX Prostate; cancer; cytostatic; gene therapy; marker; ss.
XX
XX Homo sapiens.
XX
XX MO2003009814-A2.
XX
XX 06-FEB-2003.
XX
XX 25-JUL-2002; 2002WO-US023913.
XX
XX 25-JUL-2001; 2001US-0307982P.
XX 22-AUG-2001; 2001US-0314356P.
XX 25-SEP-2001; 2001US-0325020P.
XX 12-DEC-2001; 2001US-0341746P.
XX 05-MAR-2002; 2002US-0362158P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Schlegel R, Morahan JF, Endege WO, Gannavarapu M, Gorbacheva B;
XX Hoerish S, Kamatkar S, Monsey AM, Glatt K, Zhao X, Anderson D;
XX WPI; 2003-248033/24.
XX
XX New nucleic acid molecule, useful for diagnosing or treating prostate
XX cancer.
XX
XX Disclousure; SEQ ID NO 404; 99pp; English.
XX
XX The invention relates to newly discovered cancer markers associated with
XX the cancerous state of prostate cells. Also disclosed is a method of
XX assessing whether a patient is afflicted with prostate cancer. The method
XX of the invention involves assessing whether a patient is afflicted with
XX prostate cancer by comparing the level of expression of a marker in a
XX patient sample and the normal level of expression of the marker in a
XX control non-prostate cancer sample, where a significant increase in the
XX level of expression of the marker in the patient sample and the normal
XX level indicates that the patient is afflicted with prostate cancer.
XX Nucleic acids of the invention are useful for diagnosing or treating
XX prostate cancer, and may be useful in gene therapy. Sequences given in
XX ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence
XX data for this patent did not form part of the printed specification, but
XX was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1027 BP; 307 A; 232 C; 200 G; 288 T; 0 U; 0 Other;
Query Match 100.0%; Score 1027; DB 10; Length 1027;
Best Local Similarity 100.0%; Pred. No. 4,3e-290; Indels 0; Gaps 0;
Matches 1027; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGCAAGAGTTGGGCAAAAAAATCAAGGTATTTGGTCCCGGAACAAGCTTATCATTAACA 60
DB 1 GGGCAAGAGTTGGGCAAAAAAATCAAGGTATTTGGTCCCGGAACAAGCTTATCATTAACA 60
QY 61 GATTAACAACCTTGTATGACAGATGTTTCCCGCAAGCCCATATTTTCTTCTCAATTGCT 120
DB 61 GATTAACAACCTTGTATGACAGATGTTTCCCGCAAGCCCATATTTTCTTCTCAATTGCT 120

```

[illegible]

DT 18-NOV-2004 (first entry)  
 DE Human. PRO cDNA sequence SEQ ID NO:838.  
 XX  
 XX human; PRO; immune related disease; inflammatory immune response;  
 KW immune response stimulation; antiallergic; antianemic; antiarthritic;  
 KW antiaesthetic; antidiabetic; antinflammatory; antiporiatic;  
 KW antirheumatic; antithyroid; CNS; dermatological; gastrointestinal;  
 KW haemostatic; hepatotropic; immunostimulant; immunosuppressive; muscular;  
 KW nephrotropic; neuroprotective; osteopathic; respiratory; vasotropic;  
 KW vitucide; gene therapy; gene; ss.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO2004039956-A2.  
 XX  
 XX 13-MAY-2004.  
 XX  
 XX 28-OCT-2003; 2003WO-US034381.  
 XX  
 XX 29-OCT-2002; 2002US-0422472P.  
 XX  
 XX (GENTH ) GENTECH INC.  
 XX  
 XX Aggarwal S, Clark H, Gurney AL, Schoenfeld J, Williams PM;  
 PI Wood WI, Wu TD;  
 XX  
 XX WPI; 2004-376182/35.  
 DR P-PSDB; ADP54863.  
 PT  
 PT New PRO polynucleotides and polypeptides, useful in diagnosing  
 PT and treating an immune related disease, e.g. systemic lupus  
 PT erythematosus, rheumatoid arthritis, diabetes mellitus or asthma and in  
 PT stimulating an immune response.  
 XX  
 XX  
 PS Claim 2; SEQ ID NO 838; 3009pp; English.  
 XX  
 XX The present invention describes an isolated PRO nucleic acid (1). Also  
 CC described: (1) a vector comprising (1); (2) a host cell comprising the  
 CC vector of (1); (3) a process for producing a PRO polypeptides; (4) an  
 CC isolated PRO polypeptide; (5) a chimeric molecule comprising the  
 CC polypeptide of (4) fused to a heterologous amino acid sequence; (6) an  
 CC antibody which specifically binds to a polypeptide of (4); (7) a  
 CC composition of matter comprising a polypeptide of (4), an agonist or  
 CC antagonist of the polypeptide or an antibody that binds to the  
 CC polypeptide in combination with a carrier; (8) an article of manufacture  
 CC comprising a container, a label on the container and a composition of  
 CC matter of (7); (9) a method of treating an immune related disease in a  
 CC mammal; (10) a method for determining the presence of a PRO polypeptide  
 CC in a sample suspected of having the polypeptide; (11) a method of  
 CC diagnosing an immune related disease or an inflammatory immune response  
 CC in mammal; (12) a method of identifying a compound that inhibits or  
 CC mimics the activity of or expression of a gene encoding a PRO polypeptide  
 CC; and (13) a method of stimulating the immune response in a mammal. The  
 CC PRO sequences have antiallergic, antianemic, antiarthritic,  
 CC antiaesthetic, antidiabetic, antinflammatory, antiporiatic,  
 CC antirheumatic, antithyroid, CNS, dermatological, gastrointestinal,  
 CC haemostatic, hepatotropic, immunostimulant, immunosuppressive, muscular,  
 CC nephrotropic, neuroprotective, osteopathic, respiratory, vasotropic and  
 CC vitucide activities, and can be used in gene therapy. The nucleic acid  
 CC (1) and the encoded polypeptides, compositions, kits and methods are  
 CC useful in diagnosing and treating an immune related disease and in  
 CC stimulating an immune response. The present sequence represents a human  
 CC PRO nucleotide sequence from the present invention.  
 XX  
 SO Sequence 1027 BP; 307 A; 232 C; 200 G; 288 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 1027; DB 13; Length 1027;  
 Best Local Similarity 100.0%; Prid. No. 4.3e-290;  
 Matches 1027; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 GGGCAAGGTTGGCGAATAAATCAAGTATTGTCGCGGAACAAGCTATCATTCACA 60  
 |||||||

```

Db      1 GGGCAGAGTTGGGCAAAAAAATAGGTATTGGTCCGAGACAAAGCTTATCATTTACA 60
Qy      61 GATTAACCACTTGATGAGATGTTTCCCAAGCCCACTATTTTCTTCTTCATTTGCT 120
Db      61 GATTAACCACTTGATGAGATGTTTCCCAAGCCCACTATTTTCTTCTTCATTTGCT 120
Qy      121 GAAACAAAGCTCCGAAAGGCTGGAACATCTTTGTTCTTGTGAAATTTTCCCTGAT 180
Db      121 GAAACAAAGCTCCGAAAGGCTGGAACATCTTTGTTCTTGTGAAATTTTCCCTGAT 180
Qy      181 GTTATTAAGATACATTTGGCAAGAAAGAGCAACGATTTGGGATCCGAGAGGG 240
Db      181 GTTATTAAGATACATTTGGCAAGAAAGAGCAACGATTTGGGATCCGAGAGGG 240
Qy      241 AACACATGAAGACTAAACAACATACATGAATTTAGTGTGTTACGGTGCCAGAAAG 300
Db      241 AACACATGAAGACTAAACAACATACATGAATTTAGTGTGTTACGGTGCCAGAAAG 300
Qy      301 TCACATGACAAAGAACACAGATGTATGTCAGACATGAGATTAATAAAACGAGTTGAT 360
Db      301 TCACATGACAAAGAACACAGATGTATGTCAGACATGAGATTAATAAAACGAGTTGAT 360
Qy      361 CAAGAAATTAATCTTCCCAATAAGACGGATGTCAATCAATGATCCCAAGACAT 420
Db      361 CAAGAAATTAATCTTCCCAATAAGACGGATGTCAATCAATGATCCCAAGACAT 420
Qy      421 TGTTCAAAGATGCAAAATGATACACTACTGTCAGAGCTCAAAACACTCTGCATATTAC 480
Db      421 TGTTCAAAGATGCAAAATGATACACTACTGTCAGAGCTCAAAACACTCTGCATATTAC 480
Qy      481 ATGTACTCTCTCTGCTCTCTCAAGATGTGTCTATTTTGGCATCATCTCTGCTCTG 540
Db      481 ATGTACTCTCTCTGCTCTCTCAAGATGTGTCTATTTTGGCATCATCTCTGCTCTG 540
Qy      541 CTTAGAGAAACGGCTTCTGTCGATGAGAGAAATCATTAACAGAGGTGGCAAGAGA 600
Db      541 CTTAGAGAAACGGCTTCTGTCGATGAGAGAAATCATTAACAGAGGTGGCAAGAGA 600
Qy      601 GGCCATCTTTTCTCATCGGTTATTTGTCCTTAGAAGGCTCTTGAAGATCTAATTGGGC 660
Db      601 GGCCATCTTTTCTCATCGGTTATTTGTCCTTAGAAGGCTCTTGAAGATCTAATTGGGC 660
Qy      661 TTTCTTCTGGGTTTGGGCCATTTCAAGTTCTCATGTGTGTAATTTCTATATTATTTGA 720
Db      661 TTTCTTCTGGGTTTGGGCCATTTCAAGTTCTCATGTGTGTAATTTCTATATTATTTGA 720
Qy      721 TTAAGGTTTCAAAACAGTGGGACACAGAGAACCTCATCTGTAATTAACAATGAGGAT 780
Db      721 TTAAGGTTTCAAAACAGTGGGACACAGAGAACCTCATCTGTAATTAACAATGAGGAT 780
Qy      781 AGCCACGGCGATCTCCAGACCAATCTCTCCATGTTTTCACAGCTCTCCAGCCAAACC 840
Db      781 AGCCACGGCGATCTCCAGACCAATCTCTCCATGTTTTCACAGCTCTCCAGCCAAACC 840
Qy      841 AAATAGGCGCTGCTAATGATGTAACATCTCGGGCTTTAGCCTTGTCTCTTAGTG 900
Db      841 AAATAGGCGCTGCTAATGATGTAACATCTCGGGCTTTAGCCTTGTCTCTTAGTG 900
Qy      901 TTTCTTATCAGATPACTGCTGGAAGCTTCAATTTTACAGGCCCTGGAAGCAGCTTCT 960
Db      901 TTTCTTATCAGATPACTGCTGGAAGCTTCAATTTTACAGGCCCTGGAAGCAGCTTCT 960
Qy      961 TTGCTAGTTGAATTAATGATGATGTTTTCGGTAAATAGCAAAATTAATTTAAAAAATG 1020
Db      961 TTGCTAGTTGAATTAATGATGATGTTTTCGGTAAATAGCAAAATTAATTTAAAAAATG 1020
Qy      1021 AAAAGTT 1027
Db      1021 AAAAGTT 1027

```

RESULT 4  
ADM44400

```

ID      ADM44400 standard; DNA; 1027 BP.
XX
AC      ADM44400;
XX
DT      24-MAR-2005 (first entry)
XX
DE      Human TARP polynucleotide.
XX
KW      T-cell receptor gamma alternate reading frame protein; TARP;
KW      immunogenicity; breast tumor; prostate cancer; cytostatic; neoplasm; gene;
XX      ds.
OS      Homo sapiens.
XX
FH      Key      Location/Qualifiers
FT      CDS      74..250
FT      /tag= a
FT      /product= "Human TARP"
XX
PF      WO200500889-A1.
XX
PD      06-JAN-2005.
XX
PR      02-JUN-2004; 2004MO-US017574.
XX
PR      05-JUN-2003; 2003US-0476467P.
XX
PI      (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI      Berzofsky JA, Oh S, Pastan I;
XX
XX      WPI; 2005-091494/10.
XX      DR      P-PSDB; ADM44399.
XX      DR      GENBANK; AF151103.
XX
PT      New T-cell receptor gamma alternate reading frame protein (TARP)
PT      polypeptide, useful for treating a subject having breast or prostate
PT      cancer and for generating an immune response to TARP-expressing breast
PT      and prostate cancer cells.
XX
PS      Disclosure; SEQ ID NO 2; 83pp; English.
XX
XX      The invention relates to an immunogenic T-cell receptor gamma alternate
XX      reading frame protein (TARP) polypeptide and the polynucleotide encoding
XX      it. The invention also relates to a vector comprising the TARP
XX      polynucleotide, a host cell transformed with the vector, a pharmaceutical
XX      composition comprising a therapeutic amount of the polypeptide or the
XX      polynucleotide in a pharmaceutical carrier, eliciting an immune response
XX      in a subject, inhibiting the growth of a breast cancer or a prostate
XX      cancer cell, a reagent comprising a tetramer of the polypeptide bound to
XX      HLA-A2 and streptavidin, where the reagent is labeled or unlabeled, and
XX      detecting T cells expressing CD8 that specifically recognize the TARP
XX      polypeptide in a subject. Eliciting an immune response in a subject
XX      comprises administering the polypeptide or polynucleotide, thus producing
XX      an immune response in the subject. The immune response comprises a T cell
XX      response or inducing cytotoxic T cells that induce lysis of cells
XX      expressing the TARP polypeptide. The subject has breast cancer or
XX      prostate cancer. The immune response decreases the growth of the prostate
XX      cancer or breast cancer. The method further comprises administering an
XX      adjuvant to the subject. Inhibiting the growth of a breast cancer or
XX      prostate cancer cell comprises culturing cytotoxic T lymphocytes (CTLs)
XX      or CTL precursor cells with the polypeptide and an antigen presenting
XX      cell to produce activated CTLs matured from the CTL precursors that
XX      recognize the breast cancer or the prostate cancer cells, and contacting
XX      the breast cancer or the prostate cancer cell with the activated CTLs or
XX      CTLs matured from the CTL precursors, thus inhibiting the growth of the
XX      breast cancer or the prostate cancer cell. The TARP polypeptide and
XX      polynucleotide are useful for treating a subject having breast or
XX      prostate cancer. The polypeptide is useful for generating an immune
XX      response to breast cancer and prostate cancer cells that express TARP
XX      polypeptides. This sequence represents the human TARP polynucleotide of
XX      the invention.

```

Sequence 1027 BP; 307 A; 232 C; 200 G; 288 T; 0 U; 0 Other;

Query Match 100.0%; Score 1027; DB 14; Length 1027;  
Best Local Similarity 100.0%; Pred. No. 4.3e-290;  
Matches 1027; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GGGCAAGAGTGGGCAAAAAATCAAGTATTTGGTCCCGGAACAAAGCTTATCATTAACA 60
DB 1 GGGCAAGAGTGGGCAAAAAATCAAGTATTTGGTCCCGGAACAAAGCTTATCATTAACA 60
QY 61 GATTAACAATTGATGAGATGTTTCCCAAGCCCACTATTTTCTCTCAATTGCT 120
DB 61 GATTAACAATTGATGAGATGTTTCCCAAGCCCACTATTTTCTCTCAATTGCT 120
QY 121 GAAACAAAGCTCCGAAGGCTGGAAACATCTTTGCTCTTGGAAATTTTCCCGAT 180
DB 121 GAAACAAAGCTCCGAAGGCTGGAAACATCTTTGCTCTTGGAAATTTTCCCGAT 180
QY 181 GTTATTAAGATACATTGGCAAGAAAGAGCAACAGATTCTGGATCCCAAGAGGG 240
DB 181 GTTATTAAGATACATTGGCAAGAAAGAGCAACAGATTCTGGATCCCAAGAGGG 240
QY 241 AACACATGAAGACTAAAGACATACATGAATTTAGCTGTTAACGGTCCAGAAAG 300
DB 241 AACACATGAAGACTAAAGACATACATGAATTTAGCTGTTAACGGTCCAGAAAG 300
QY 301 TCATTCGACAAAGAAACACAGATGATGCTCGACATGAGATTAATAAAACGAGTTGAT 360
DB 301 TCATTCGACAAAGAAACACAGATGATGCTCGACATGAGATTAATAAAACGAGTTGAT 360
QY 361 CAAGAAATTTATCTTCTCCCAATTAAGACGAGTGCATCAATGATGCCAAGACAT 420
DB 361 CAAGAAATTTATCTTCTCCCAATTAAGACGAGTGCATCAATGATGCCAAGACAT 420
QY 421 TGTTCAAAAGATGCAAAATGATACACTCTGCTGACGTCAACAAACCTCTGATATTAC 480
DB 421 TGTTCAAAAGATGCAAAATGATACACTCTGCTGACGTCAACAAACCTCTGATATTAC 480
QY 481 ATGTAACCTCTCTGCTGCTCAAGAGTGTGATTTTGGCATCATCACTGCTGCTG 540
DB 481 ATGTAACCTCTCTGCTGCTCAAGAGTGTGATTTTGGCATCATCACTGCTGCTG 540
QY 541 CTTAGAAAGAACGGCTTCTGCTGAATGAGAGAAATCATACAGACGGTGGCAAGAGA 600
DB 541 CTTAGAAAGAACGGCTTCTGCTGAATGAGAGAAATCATACAGACGGTGGCAAGAGA 600
QY 601 GGCCATCTTTTCTCATCGATTATTTGTCCTAGAAAGCTCTTCTGAGGATCTAGTTGGGC 660
DB 601 GGCCATCTTTTCTCATCGATTATTTGTCCTAGAAAGCTCTTCTGAGGATCTAGTTGGGC 660
QY 661 TTTCTTTCTGGGTTGGGCCATTTCAAGTTCATGATGTAATTTCTATCATTTATTTGTA 720
DB 661 TTTCTTTCTGGGTTGGGCCATTTCAAGTTCATGATGTAATTTCTATCATTTATTTGTA 720
QY 721 TTAAGGTTTAAACCAAGTGGGACACAGAGAACCTCACTGTAATTAACAATGAGGAT 780
DB 721 TTAAGGTTTAAACCAAGTGGGACACAGAGAACCTCACTGTAATTAACAATGAGGAT 780
QY 781 AGCCACGGCGATCTCTCAAGCAAAATCTCTCATGTTTTCACAGCTCTCTCAAGCAACC 840
DB 781 AGCCACGGCGATCTCTCAAGCAAAATCTCTCATGTTTTCACAGCTCTCTCAAGCAACC 840
QY 841 AAATAGCGCTGCTAATGATGTAACATCTCGGGCTTACGCTTGTCTCTCTAGTG 900
DB 841 AAATAGCGCTGCTAATGATGTAACATCTCGGGCTTACGCTTGTCTCTCTAGTG 900
QY 901 TTTCTTATTCAGATACGCTGGAAGCTTCAATTTTACAGGCTCGAAGACAGCTTCT 960
DB 901 TTTCTTATTCAGATACGCTGGAAGCTTCAATTTTACAGGCTCGAAGACAGCTTCT 960
QY 961 TTGCTAGTTGAATTAATGATGATGTTTTCGTATTAAGCAAAATTAATTTAAAAAATG 1020
DB 961 TTGCTAGTTGAATTAATGATGATGTTTTCGTATTAAGCAAAATTAATTTAAAAAATG 1020

```

QY 1021 AAAAGTT 1027  
DB 1021 AAAAGTT 1027

RESULT 5  
ADL06443  
ID ADL06443 standard; cDNA; 1395 BP.

AC ADL06443;  
XX  
DT 20-MAY-2004 (first entry)  
XX

DE Human tumour-associated antigenic target (TAT) cDNA sequence #23.

XX Human; tumour-associated antigenic target; TAT; cell death; tumour;  
XX cancer; cytostatic; gene; ss.

XX Homo sapiens.

XX W02004016225-A2.

XX 26-FEB-2004.

XX 19-AUG-2003; 2003WO-US025892.

XX 19-AUG-2002; 2002US-0404809P.

XX 21-AUG-2002; 2002US-0405645P.

XX 23-SEP-2002; 2002US-0413192P.

XX 15-OCT-2002; 2002US-041908P.

XX 15-NOV-2002; 2002US-0426847P.

XX 02-JUL-2003; 2003US-0484959P.

XX (GENT) GENENTECH INC.

XX Desauvage FJ, Frantz G, Hillan KJ, Polakis P, Polson A, Smith V;

XX Spencer SD, Wu TD, Zhang Z;

XX WPI; 2004-257144/24.

XX DR P-PSDB; ADL06523.

XX PT New antibody that binds to a tumor-associated antigenic target (TAT)

XX PT polypeptide, useful for preparing a composition for diagnosing or

XX PT treating cancer.

XX PS Claim 1; SEQ ID NO 23; 319pp; English.

XX CC The present invention relates to the isolation of human tumour-associated

XX CC antigenic target (TAT) polynucleotide and polypeptide sequences. Also

XX CC disclosed is an antibody that binds to a TAT polypeptide. The antibody is

XX CC a monoclonal antibody, an antibody fragment, a chimeric antibody or a

XX CC humanised antibody. It is conjugated to a growth inhibitory agent. It is

XX CC produced in bacteria or in CHO cells and induces death of a cell to which

XX CC it binds. The antibody is useful for preparing a composition for

XX CC diagnosing or treating tumours and cancer. The present sequence

XX CC represents a human TAT cDNA sequence of the invention.

Sequence 1395 BP; 405 A; 324 C; 286 G; 380 T; 0 U; 0 Other;

Query Match 98.7%; Score 1013.8; DB 12; Length 1395;  
Best Local Similarity 99.8%; Pred. No. 3.7e-286;  
Matches 1015; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 4 CAAGAGTTGGGCAAAAAATCAAGTATTTGGTCCCGGAACAAAGCTTATCATTAACAGAT 63
DB 379 CAAGAGTTGGGCAAAAAATCAAGTATTTGGTCCCGGAACAAAGCTTATCATTAACAGAT 438
QY 64 AAACAATTGATGAGATGTTTCCCAAGCCCACTATTTTCTCTCTCAATTGCTGAA 123
DB 439 AAACAATTGATGAGATGTTTCCCAAGCCCACTATTTTCTCTCTCAATTGCTGAA 498
QY 124 ACAAGCTCAGAAAGCTGGAACATACCTTGTCTTGTGAGAAATTTTCCCTGATGTT 183

```

```
Db 499 ACAAGCTCCAGAAAGCTGGAACATCTTTGCTCTTGGAAATTTTCCGAGTT 558
Qy 184 ATTAAGATACATTTGGCAAGAAAGAGCAACAGATTTGGGATCCGAGAGGCAAC 243
Db 559 ATTAAAGATACATTTGGCAAGAAAGAGCAACAGATTTGGGATCCGAGAGGCAAC 618
Qy 244 ACCATGAAGACTAAAGACATACATGAAATTTAGCTGTTAACGGTGCCAGAAAGTCA 303
Db 619 ACCATGAAGACTAAAGACATACATGAAATTTAGCTGTTAACGGTGCCAGAAAGTCA 678
Qy 304 CTGACAAAGAAACACAGATGATGTCAGACATGAGATTAATAAAGAGGATGATTA 363
Db 679 CTGACAAAGAAACACAGATGATGTCAGACATGAGATTAATAAAGAGGATGATTA 738
Qy 364 GAAATTTATCTTTCTCCATTAAGACGAGATGTCATCAATGATGCCAAAGACATTTGT 423
Db 739 GAAATTTATCTTTCTCCATTAAGACGAGATGTCATCAATGATGCCAAAGACATTTGT 798
Qy 424 TCAAAAGATGCAAAATGATACACTACTGCTGACAGTCAAAACACCTCTGCATATTACATG 483
Db 799 TCAAAAGATGCAAAATGATACACTACTGCTGACAGTCAAAACACCTCTGCATATTACACG 858
Qy 484 TACCTCCCTGCTGCTCCTCAAGAGTGGTCTATTGTCATCATCACTGCTGCTGCTT 543
Db 859 TACCTCCCTGCTGCTCCTCAAGAGTGGTCTATTGTCATCATCACTGCTGCTGCTT 918
Qy 544 AGAAGAACGGCTTTCTGCTGCAATGAGAGAAATCATTAACAGACGGTGGCAAGAGAGC 603
Db 919 AGAAGAACGGCTTTCTGCTGCAATGAGAGAAATCATTAACAGACGGTGGCAAGAGAGC 978
Qy 604 CATCTTTCTCTCATCGGTTATTTGCTCCAGAAAGCGTCTTCTGAGAGATCAAGTGGGCTTT 663
Db 979 CATCTTTCTCTCATCGGTTATTTGCTCCAGAAAGCGTCTTCTGAGAGATCAAGTGGGCTTT 1038
Qy 664 CTCTCTGGGCTTTGGGCAATTTCAAGTTCATGTCATGTCATTTCTCATATATTATTTAA 723
Db 1039 CTCTCTGGGCTTTGGGCAATTTCAAGTTCATGTCATGTCATTTCTCATATATTATTA 1098
Qy 724 CGGTTTCAAAACAGATGGGACACAGAGAACCTCACTCTGTAATTAACAATGAGAAATAGC 783
Db 1099 CGGTTTCAAAACAGATGGGACACAGAGAACCTCACTCTGTAATTAACAATGAGAAATAGC 1158
Qy 784 CACGGCGATCTCCAGACCAATCTCTCATGTTTTCACAGCTCTCTCCAGCCAAACCCAA 843
Db 1159 CACGGCGATCTCCAGACCAATCTCTCATGTTTTCACAGCTCTCTCCAGCCAAACCCAA 1218
Qy 844 TAGCGCCGCTAATAGTGAACATCTCGCGCTTCTAGCCCTTCTTATGTTTC 903
Db 1219 TAGCGCCGCTAATAGTGAACATCTCGCGCTTCTAGCCCTTCTTATGTTTC 1278
Qy 904 TTTAATCAGATPACTGCTGGAAGCTTTCAATTTTACAGCCCTGGAACAGATCTTTTGG 963
Db 1279 TTTAATCAGATPACTGCTGGAAGCTTTCAATTTTACAGCCCTGGAACAGATCTTTTGG 1338
Qy 964 CTAGTTGAATTAATGCTGCTGTTTTCGGTAATAAGCAAAATTAATTTAAAAAATG 1020
Db 1339 CTAGTTGAATTAATGCTGCTGTTTTCGGTAATAAGCAAAATTAATTTAAAAAATG 1395

RESULT 6
ADY16296 ID ADY16296 standard; DNA; 1530 BP.
XX AC ADY16296;
XX 05-MAY-2005 (first entry)
XX DNA encoding a PRO polypeptide, SEQ ID NO 2102.
KW Antinflammatory; Immune disorder; Dermatological; Immunosuppressive;
KW Antirheumatic; Antiarthritic; Osteopathic; Hemostatic; Antianemic;
KW Antihypoid; Antidiabetic; Nephroretropic; CNS-Gen.; Hepatotropic;
```

```
KW Virucide; Gastrointestinal-Gen.; Antipeoriatic; Antialsthmatic;
KW Antiallergic; de; gene; diagnosis.
XX OS Homo sapiens.
XX PN MO2005016962-A2.
XX 24-FEB-2005.
XX 11-AUG-2004; 2004MO-US026249.
XX 11-AUG-2003; 2003US-0493546P.
XX (GERTH ) GENENTECH INC.
PI Abbas A, Clark H, Ouyang W, Williams MP, Wood WI, Wu TD;
DR WPI; 2005-182330/19.
XX New nucleic acid encoding PRO polypeptide, useful for diagnosing and
PT treating an immune related disorder, e.g. systemic lupus erythematosus,
PT rheumatoid arthritis, osteoarthritis, thyroiditis, or diabetes mellitus.
XX Claim 1; SEQ ID NO 2102; 158pp; English.
XX The invention relates to an isolated nucleic acid encoding a PRO
XX polypeptide. The polypeptide, agonist or an antagonist, antibody,
XX composition, and method are useful for diagnosing and treating an immune
XX related disorder, e.g. systemic lupus erythematosus, rheumatoid
XX arthritis. The present sequence represents a DNA encoding a PRO
XX polypeptide.
XX SQ Sequence 1530 BP; 435 A; 361 C; 322 G; 412 T; 0 U; 0 Other;
Query Match 95.8%; Score 984; DB 14; Length 1530;
Best Local Similarity 98.0%; Pred. No. 2,1e-277;
Matches 996; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
Qy 12 GGGCAAAAATCAAGGATTTGTCGCCGAAACCAAGCTTATCATTAACATTAACAACCT 71
Db 515 GGGAAATTAATAGAAACTCTTGACAGTGGAAACAACCTTGTCACAGATTAACAACCT 574
Qy 72 TGATGCAAGATTTTCCCCCAAGCCCACTATTTTCTTCTTCAATTTGCTGAAACAAAGCT 131
Db 575 TGATGCAAGATTTTCCCCCAAGCCCACTATTTTCTTCTTCAATTTGCTGAAACAAAGCT 634
Qy 132 CCAAGAGGCTGGAACATACCTTTGTCCTTGAAGAAATTTTCCCTGATGTTATTAAGAT 191
Db 635 CCAAGAGGCTGGAACATACCTTTGTCCTTGAAGAAATTTTCCCTGATGTTATTAAGAT 694
Qy 192 ACATTGGCAAGAAAGAAAGACCAACAGATTTCTGGGATCCGAGAGGAGAAACACCATGAA 251
Db 695 ACATTGGCAAGAAAGAAAGACCAACAGATTTCTGGGATCCGAGAGGAGAAACACCATGAA 754
Qy 252 GACTTAAGACATACATGAAATTTAGTGGTTAACGGTGCAGAAAAAGTACCTGAGCAA 311
Db 755 GACTTAAGACATACATGAAATTTAGTGGTTAACGGTGCAGAAAAAGTACCTGAGCAA 814
Qy 312 AGAACACAGATGTAATCGTCAGACATGAGAAATTAATAAAGAGAGTGTGATCAAGAAATTA 371
Db 815 AGAACACAGATGTAATCGTCAGACATGAGAAATTAATAAAGAGAGTGTGATCAAGAAATTA 874
Qy 372 CTCTCTCCATTAAGACGAGATGTCATCAATGAGATCCCAAGACAAATTTGTCAAAAAGA 431
Db 875 CTCTCTCCATTAAGACGAGATGTCATCAATGAGATCCCAAGACAAATTTGTCAAAAAGA 934
Qy 432 TGCAAAATGATACACTACTGCTGACGCTCAAAACACCTCTGCAATTAACATGATCTCT 491
Db 935 TGCAAAATGATACACTACTGCTGACGCTCAAAACACCTCTGCAATTAACATGATCTCT 994
Qy 492 CTTGCTCTCTCAAGAGTGGTGTCTATTTTGGCATCATCAACCGTGTGCTTAGAAGAG 551
Db 995 CTTGCTCTCTCAAGAGTGGTGTCTATTTTGGCATCATCAACCGTGTGTCTTAGAAGAG 1054
```

QY 552 GGGTTTGTGTCGCAATGAGAGAAATCATTAACAGACGGTGACCAAGAGGCCATCTTTT 611  
DB 1055 GGGTTTGTGTCGCAATGAGAGAAATCATTAACAGACGGTGACCAAGAGGCCATCTTTT 1114  
QY 612 CCTCATCGGTTATGTCCTTGAAGACGCTCTTGAGAGATCTAGTTGGGCTTTCTTTCTGG 671  
DB 1115 CCTCATCGGTTATGTCCTTGAAGACGCTCTTGAGAGATCTAGTTGGGCTTTCTTTCTGG 1174  
QY 672 GTTTGGGCAATTTCAAGTTCTCATGTGTACTATTTCTATTAATTTGTAACGGTTTTC 731  
DB 1175 GTTTGGGCAATTTCAAGTTCTCATGTGTACTATTTCTATTAATTTGTAACGGTTTTC 1234  
QY 732 AAACCGATGGGACACAGAGAACCTCCTGTAATAACATGAGATAGCCAGCGA 791  
DB 1235 AAACCGATGGGACACAGAGAACCTCCTGTAATAACATGAGATAGCCAGCGA 1294  
QY 792 TCTCAGACCAATCTCTCCATGTTTTCACAGCTCCTCCAGCCAAACCAATAGCCCT 851  
DB 1295 TCTCAGACCAATCTCTCCATGTTTTCACAGCTCCTCCAGCCAAACCAATAGCCCT 1354  
QY 852 GCTATAGTGAACATCTCGGCGCTTCTAGCCTTGTCCCTCTTATGTTCTTTAATCA 911  
DB 1355 GCTATAGTGAACATCTCGGCGCTTCTAGCCTTGTCCCTCTTATGTTCTTTAATCA 1414  
QY 912 GATAACCTGCGGAGAGCTTTCATTTTACAGCCCTGAAGAGCTTCTTGTAGTTGA 971  
DB 1415 GATAACCTGCGGAGAGCTTTCATTTTACAGCCCTGAAGAGCTTCTTGTAGTTGA 1474  
QY 972 ATTATGTGTGTGTTTCCGTAAATGAACAAATTTTAAAAAATGAAAAGTT 1027  
DB 1475 ATTATGTGTGTGTTTCCGTAAATGAACAAATTTTAAAAAATGAAAAGTT 1530  
RESULT 7  
ID AB235381  
AB235381 standard; cDNA; 1799 BP.  
XX  
AC AB235381;  
XX  
DT 05-FEB-2003 (first entry)  
XX  
DE Human gene expression profile polynucleotide SEQ ID NO 492.  
XX  
KM Human; artery; endothelium; umbilical; vein; aorta; pulmonary artery;  
KM bronchial epithelium; prostate; muscle; lung fibroblast; osteoblast;  
KM tumour; microarray; genome mapping; antibiotic; antiviral; antifungal;  
KM gene expression; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN W0200274979-A2.  
XX  
PD 26-SEP-2002.  
XX  
PF 20-MAR-2002; 2002W0-US008456.  
XX  
PR 20-MAR-2001; 2001US-0276947P.  
XX  
PA (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.  
XX  
PI Wan J, Wang Y;  
XX  
DR WPI; 2002-740862/80.  
XX  
PT New gene expression profile generated from primary, endothelial,  
PT epithelial, and muscle cell types, useful for identifying disease  
PT pathologies involving alterations of gene expression, e.g. cancer.  
XX  
PS Example 3; Page 634-635; 850bp; English.  
XX  
CC The invention relates to a gene expression profile comprising one or more  
CC genes (AB234889-AB235692) and generated from a cell type. The cell type

CC is a coronary artery endothelium, umbilical artery or vein endothelium,  
CC aortic endothelium, dermal microvascular endothelium, pulmonary artery  
CC endothelium, myometrium microvascular endothelium, keratinocyte  
CC epithelium, bronchial epithelium, mammary epithelium, prostate  
CC epithelium, renal cortical epithelium, renal proximal tubule epithelium,  
CC small airway epithelium, renal epithelium, umbilical artery smooth  
CC muscle, neonatal dermal fibroblast, pulmonary artery smooth muscle,  
CC dermal fibroblast, neural progenitor cells, skeletal muscle, astrocytes,  
CC aortic smooth muscle, mesangial cells, coronary artery smooth muscle,  
CC bronchial smooth muscle, uterine smooth muscle, lung fibroblast,  
CC osteoblasts or prostate stromal cell. The gene expression profile is used  
CC for determining the level of RNA expression for a sample, determining the  
CC phenotype of a cell and distinguishing cell types. The gene or a protein  
CC expression profile is useful in identifying disease pathologies involving  
CC alterations of gene expression. The assessment of expression profiles may  
CC provide meaningful information with respect to tumor type and stage.  
CC treatment methods, and prognosis. The gene or protein expression profile  
CC may also be used for creating microarrays. The microarray is useful for  
CC genetic and physical mapping of genomes, DNA sequencing, genetic or  
CC medical diagnosis, genotyping of organisms, confirming cell or tissue  
CC identifications and in identifying promising antibiotics, antiviral or  
CC antifungal agents

XX  
SQ Sequence 1799 BP; 499 A; 434 C; 396 G; 470 T; 0 U; 0 Other;

Query Match 94.3%; Score 968.2; DB 6; Length 1799;

Best Local Similarity 98.1%; Pred. No. 9,7e-273; Matches 990; Conservative 0; Mismatches 18; Indels 1; Gaps 1;

DB 20 AATCAAGTATTGTGTCGCCGGAACAAGCTTATCATTAAGATTAACAATTGATGAC 79  
DB 791 ATAGAAAGCTTTGGGAGAGGAACAACCTTGTGTACAGATTAACAACCTTGATGAC 850  
QY 80 ATGTTTCCCCCAAGCCCAATTTTCTTCTTCATTTGCTGAACAAAGCTCCAGAAG 139  
DB 851 ATGTTTCCCCCAAGCCCAATTTTCTTCTTCATTTGCTGAACAAAGCTCCAGAAG 910  
QY 140 CTGGAACATACCTTTGCTTCTTGAAGAAATTTTCCCTGATTTTATTAATGATGATG 199  
DB 911 CTGGAACATACCTTTGCTTCTTGAAGAAATTTTCCCTGATTTTATTAATGATGATG 970  
QY 200 AAGAAAAGAGAACCAACAGATTTCT-GGATCCAGAGGGGAACAACATGAAGACTTAC 258  
DB 971 AAGAAAAGAGAACCAACAGATTTCTGGGATCCAGAGGGGAACAACATGAAGACTTAC 1030  
QY 259 GACACATACATGAATTTAGCTGTTTACCGTCCAGAAAAGTCACTGCAAAAGAAC 318  
DB 1031 GACACATACATGAATTTAGCTGTTTACCGTCCAGAAAAGTCACTGCAAAAGAAC 1090  
QY 319 AGATGATATCTGTCAGACATGGAATTAATAAAGGAGTGAATCAAGAAATTTCTTCT 378  
DB 1091 AGATGATATCTGTCAGACATGGAATTAATAAAGGAGTGAATCAAGAAATTTCTTCT 1150  
QY 379 CCAATTAAGAGGATGTCATCAATGATGCCAAAGACAAATTTGTTCAAAAGATGCAAT 438  
DB 1151 CCAATTAAGAGGATGTCATCAATGATGCCAAAGACAAATTTGTTCAAAAGATGCAAT 1210  
QY 439 GATACACTACTGTCGAGCTCACAAACACTGTGATATTATCATGTACTCTCTCTC 498  
DB 1211 GATACACTACTGTCGAGCTCACAAACACTGTGATATTATCAACGTACTCTCTCTC 1270  
QY 499 CTCAAGAGTGTGTTTATTTTGGCATCATCACTGCTGTGCTTGAAGAAAGGCTTTC 558  
DB 1271 CTCAAGAGTGTGTTTATTTTGGCATCATCACTGCTGTGCTTGAAGAAAGGCTTTC 1330  
QY 559 TGCTGCAATGAGAGAAATCATTAACAGACGGTGACCAAGAGGCCATCTTTCTCATC 618  
DB 1331 TGCTGCAATGAGAGAAATCATTAACAGACGGTGACCAAGAGGCCATCTTTCTCATC 1390  
QY 619 GGTATTTGTCCTTGAAGAGCTCTTCTGAGGATCTAGTTGGGCTTTCTTCTGGTTTGG 678  
DB 1391 GGTATTTGTCCTTGAAGAGCTCTTCTGAGGATCTAGTTGGGCTTTCTTCTGGTTTGG 1450

Qy 679 CCATTTTCAGTTCTCATGCTGTGTAATCTTCTATCATTTATTTGTAATACGGTTTCAACACG 738  
Db 1451 CCATTTTCAGTTCTCATGCTGTGTAATCTTCTATCATTTATTTGTAATACGGTTTCAACACG 1510  
Qy 739 TGGGACACAGAGAACTCAGCTGTGTAATTAACAATGAGAAATAGCAGGAGATCTCCAG 798  
Db 1511 TGGGACACAGAGAACTCAGCTGTGTAATTAACAATGAGAAATAGCAGGAGATCTCCAG 1570  
Qy 799 CACCAATCTCTCATGTTTTCACAGCTCTTCAGCCCAACCAATAGCGCTGTATAG 858  
Db 1571 CACCAATCTCTCATGTTTTCACAGCTCTTCAGCCCAACCAATAGCGCTGTATAG 1630  
Qy 859 TGTAGACATCTCTGCGGCTTCTAGCCTTGTCCCTCTTCTAGTGTCTTTAATCAGTAACT 918  
Db 1631 TGTAGACATCTCTGCGGCTTCTAGCCTTGTCCCTCTTCTAGTGTCTTTAATCAGTAACT 1690  
Qy 919 GCCTGGAAGCCTTTCATTTTACAGCCCTGGAAGAGCTCTTCTGCTAGTTGAATTATGT 978  
Db 1691 GCCTGGAAGCCTTTCATTTTACAGCCCTGGAAGAGCTCTTCTGCTAGTTGAATTATGT 1750  
Qy 979 GGTGTGTTTTCCGTAAATAGCAAAATAAATTTAAAAAATGAAAAGTT 1027  
Db 1751 GGTGTGTTTTCCGTAAATAGCAAAATAAATTTAAAAAATGAAAAGTT 1799

## RESULT 8

AAD38828  
ID AAD38828 standard; cDNA; 1418 BP.

XX AC AAD38828;

XX DT 23-SEP-2002 (first entry)

XX DE Human PSNA cDNA #8.

XX KW Human; prostate specific protein; PSP; prostate specific nucleic acid;

XX KW vaccine; transgenic; prostate cancer; gene therapy; transgenic animal;

XX KW PSNA; ss.

XX OS Homo sapiens.

XX PN MO200224718-A1.

XX PD 28-MAR-2002.

XX PF 19-SEP-2001; 2001MO-US029386.

XX PR 19-SEP-2000; 2000US-0233746P.

XX PA (DIAD-) DIADEXUS INC.

XX PI Sun Y, Recipon H, Caffery R, Ali S;

XX DR WPI; 2002-471216/50.

XX PT Novel isolated prostate specific polypeptide useful for identifying,

XX PT diagnosing, monitoring, staging, imaging, and treating prostate cancer

XX PT and non-cancerous disease states in prostate.

XX PS Claim 1; Page 187; 210pp; English.

XX CC The invention relates to prostate specific proteins (PSP) and prostate

XX CC specific nucleic acids (PSNA). Sequences of the invention are useful for

XX CC identifying, diagnosing, monitoring, staging, imaging and treating

XX CC prostate cancer and non-cancerous disease states in prostate. They are

XX CC also useful for producing engineered prostate tissue for treatment and

XX CC research. The PSNA sequences are used in gene therapy and for producing

XX CC transgenic animals and cells. The invention is also used as vaccines. The

XX CC present sequence is human PSNA cDNA

XX SQ Sequence 1418 BP; 418 A; 335 C; 285 G; 380 T; 0 U; 0 Other;

Query Match 92.1%; Score 945.6; DB 6; Length 1418;

Best Local Similarity 97.8%; Pred. No. 3,6e-266;  
Matches 1001; Conservative 0; Mismatches 19; Indels 4; Gaps 4;

Qy 4 CAAGAGTTGGGCAAAAAATCAAGTATTTGGTCCCGAAACAAGTTATCATTAACAGAT 63  
Db 379 CAAGAGTTGGGCAAAAAATCAAGTATTTGGTCCCGAAACAAGTTATCATTAACAGAT 438  
Qy 64 AAACAATTGATGAGATGTTTCCCGAAGCCAGCTATTTTCTCTCTCAATTGCTGAA 123  
Db 439 AAACAATTGATGAGATGTTTCCCGAAGCCAGCTATTTTCTCTCTCAATTGCTGAA 498  
Qy 124 ACAAGCTCAGAAAGGCTGGAACATACCTTGTCTCTGGAATTTTCCCGATGTT 183  
Db 499 ACAAGCTCAGAAAGGCTGGAACATACCTTGTCTCTGGAATTTTCCCGATGTT 558  
Qy 184 ATTAAGATACATTGGCAAGAAAGAAAGCAACAAGATTCGAGATCCAGAGAGGGAAC 243  
Db 559 ATTAAGATACATTGGCAAGAAAGAAAGCAACAAGATTCGAGATCCAGAGAGGGAAC 618  
Qy 244 ACATGAAAGCTAAAGCAACATACATGAATTTAGTGTGTTAACGGTGCAGAAAAGTCA 303  
Db 619 ACATGAAAGCTAAAGCAACATACATGAATTTAGTGTGTTAACGGTGCAGAAAAGTCA 678  
Qy 304 CTGACAAAGAAACACATGATGTCAGACATGAAATTAATAACGAGTTGATCAA 363  
Db 679 CTGACAAAGAAACACATGATGTCAGACATGAAATTAATAACGAGTTGATCAA 738  
Qy 364 GAAATTTATCTTCTCCCAATTAAGACGAGATGTCATCAATGATGCCAAGACAAATTGT 423  
Db 739 GAAATTTATCTTCTCCCAATTAAGACAGATGTCATCAATGATGCCAAGACAAATTGT 798  
Qy 424 TCAAAAGATGCAAAATGATACCTACTGCTGACGCTCAACAAACCTCTGCAATTAACATG 483  
Db 799 TCAAAAGATGCAAAATGATACCTACTGCTGACGCTCAACAAACCTCTGCAATTAACATG 858  
Qy 484 TACCTCTCTCTGCTCCCTCAAGAGTGTGTCATTTGGCAATCACTGCTGCTGTT 543  
Db 859 TACCTCTCTCTGCTCCCTCAAGAGTGTGTCATTTGGCAATCACTGCTGCTGTT 918  
Qy 544 AGAAGAAAGGCTTCTGCTGCAATGAGAGAAATCATTAACAGAGGTGGCAAGAGAGGC 603  
Db 919 AGAAGAAAGGCTTCTGCTGCAATGAGAGAAATCATTAACAGAGGTGGCAAGAGAGGC 978  
Qy 604 CATCTTTTCTCTC-ATGCGTATTTGCTCTAGAACGCTTCTGAGATTAATGTTGGCTT 662  
Db 979 CATCTTTTCTCTCAATCGGTATTTGCTCTAGAACGCTTCTGAGATTAATGTTGGCTT 1038  
Qy 663 TCTTTCTGGGTTGGGCAATTCAGTTCTCATGTTGTAAT-CTATCATTAATGTA- 720  
Db 1039 TCTTTCTGGGTTGGGCAATTCAGTTCTCATGTTGTAAT-CTATCATTAATGTAAT 1098  
Qy 721 TAAAGTTTCAAAACAGTGGGCAACAGAGAACTCCTGTAATTAACATGAGAT 780  
Db 1099 TAAAGTTTCAAAACAGTGGGCAACAGAGAACTCCTGTAATTAACATGAGAT 1158  
Qy 781 AGCCACGGCGATCTCCAGACCAATCTCTCATGTTTCCACAGCTCTCCAGCAACCC 840  
Db 1159 AGCCACGGCGATCTCCAGACCAATCTCTCATGTTTCCACAGCTCTCCAGCAACCC 1218  
Qy 841 AAATAGGCGCTGCTATATGTAAGACATCTCGGCTTACCTTGTGCTCTCTATAGT 900  
Db 1219 AAATAGGCGCTGCTATATGTAAGACATCTCGC-CCCTCTACCTTGTGCTCTCTATAGT 1277  
Qy 901 TTCTTTATCAGATTAAGCTGGAAGCTTCAATTTTACAGCCCTGGAAGCAAGCTTCT 960  
Db 1278 TTCTTTATCAGATTAAGCTGGAAGCTTCAATTTTACAGCCCTGGAAGCAAGCTTCT 1337  
Qy 961 TTGCTAGTTGAATTAATGTTGTTTTCGTAATAAGCAAAATTAATTTAAAAAATG 1020  
Db 1338 CTGCTAGTTGAATTAATGTTGTTTTCGTAATAAGCAAAATTAATTTAAAAAAT 1397  
Qy 1021 AAAA 1024  
|||

Db 1398 GAA 1401

RESULT 9

ADBE08798/C

ID ADE08798 standard; DNA; 2658 BP.

XX ADE08798;

XX

DT 29-JAN-2004 (first entry)

XX

DE Novel DNA-related contig nucleotide sequence #42.

XX

KM novel gene; novel protein; tissue marker; molecular weight marker;

XX chromosome marker; genetic disorder; contig; ds.

XX

OS Unidentified.

XX

PN MO2003054152-A2.

XX

PD 03-JUL-2003.

XX

PF 10-DEC-2002; 2002WO-US039555.

XX

PR 10-DEC-2001; 2001US-0339739P.

XX

PR 11-DEC-2001; 2001US-0339453P.

XX

PR 14-MAR-2002; 2002US-0365091P.

XX

PR 14-MAR-2002; 2002US-0365384P.

XX

PR 12-APR-2002; 2002US-037281P.

XX

PR 12-APR-2002; 2002US-0372615P.

XX

PR 22-APR-2002; 2002US-0012855P.

XX

PR 24-APR-2002; 2002US-0376045P.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QH, Wang J;

PI Ghosh M, Xue AJ, Wehman T, Weng G, Zhou P, Drmanac RT, Wang Z;

PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ,

XX

DR WPI; 2003-569235/53.

XX

PT New polynucleotides, useful for expressing recombinant proteins for

PT analysis, characterization or therapeutic use, or as markers for tissues

PT in which the corresponding protein is preferentially expressed.

XX

PS Disclosure; SEQ ID NO 1864; 1177bp; English.

XX

CC The invention comprises the amino acid and coding sequences of novel

CC proteins. The DNA and protein sequences of the invention are useful as:

CC markers for tissues in which the corresponding protein is preferentially

CC expressed; as molecular weight markers on gels; as chromosome markers or

CC tags; to identify chromosomes or to map related gene positions; and to

CC compare with endogenous DNA sequences in patients to identify potential

CC genetic disorders. The present DNA sequence was used in the

CC exemplification of the invention.

XX

SQ Sequence 2658 BP; 745 A; 585 C; 578 G; 750 T; 0 U; 0 Other;

Query Match 89.1%; Score 915; DB 10; Length 2658;

Best Local Similarity 97.5%; Pred. No. 4.5e-257;

Matches 927; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 59 CAGATAAACAAGCTGATGAGATGTTTCCCAAGCCCACTATTCTTCTTCAATTG 118

DB 2396 CTGATAAACAAGCTGATGAGATGTTTCCCAAGCCCACTATTCTTCTTCAATTG 2337

QY 119 CTGAAAACAAGCTCAGAGGCTGGAACATACCTTTGCTTCTGAGAAATTTTCCCTG 178

DB 2336 CTGAAAACAAGCTCAGAGGCTGGAACATACCTTTGCTTCTGAGAAATTTTCCCTG 2277

QY 179 ATGTTATTAAGATACATTTGGCAAGAAAAGAGCAACAGCATTTCTGGATCCAGAGG 238

DB 2276 ATGTTATTAAGATACATTTGGCAAGAAAAGAGCAACAGCATTTCTGGATCCAGAGG 2217

QY 239 GGAACACCATGAGACTTAACGACACATACATGAATTTAGCTGTTAACGGTCAGAAA 298

DB 2216 GGAACACCATGAGACTTAACGACACATACATGAATTTAGCTGTTAACGGTCAGAAA 2157

QY 299 AGTCACTGACAAAGAACACAGATGTATGCTGACAGCATGAGAAATTAATAAAACGGAGTTG 358

DB 2156 AGTCACTGACAAAGAACACAGATGTATGCTGACAGCATGAGAAATTAATAAAACGGAGTTG 2097

QY 359 ATCAAGAAATTAATCTTCTCCCAATTAAGACGATGTATCAACAAATGATCCCAAGACA 418

DB 2096 ATCAAGAAATTAATCTTCTCCCAATTAAGACGATGTATCAACAAATGATCCCAAGACA 2037

QY 419 ATGTTCAAAAGATGAGAAATGATACACTACTGTGACGCTCACAAACACCTGTGATATT 478

DB 2036 ATGTTCAAAAGATGAGAAATGATACACTACTGTGACGCTCACAAACACCTGTGATATT 1977

QY 479 ACATGTACCTCTCTCTGCTCTCAAGAGTGTGCTAATTTTGGCATCATCAGCTGTCTGTC 538

DB 1976 ACATGTACCTCTCTCTGCTCTCAAGAGTGTGCTAATTTTGGCATCATCAGCTGTCTGTC 1917

QY 539 TGCTTGAGAGAACGCTTTTCTGCTGCAATGAGAGAAATCATTAACAGCGGTGACACAG 598

DB 1916 TGCTTGAGAGAACGCTTTTCTGCTGCAATGAGAGAAATCATTAACAGCGGTGACACAG 1857

QY 599 GAGGCCATCTTTCCATCATGCTTATTTGCTCCAGAAAGCTCTTGAGAGATCAGTTTG 658

DB 1856 GAGGCCATCTTTCCATCATGCTTATTTGCTCCAGAAAGCTCTTGAGAGATCAGTTTG 1797

QY 659 GCTTCTCTTCTGAGGTTGGCCATTTCACTTCACTGATGATCAATTTATCATATTATG 718

DB 1796 GCTTCTCTTCTGAGGTTGGCCATTTCACTTCACTGATGATCAATTTATCATATTATG 1737

QY 719 TATTAAGCTTTTCAAAACCACTGGGCAACAGAGAACTCACTGTATTAACATGAGA 778

DB 1736 TATTAAGCTTTTCAAAACCACTGGGCAACAGAGAACTCACTGTATTAACATGAGA 1677

QY 779 ATAGCCACGGCGATCTCCAGACCAATCTCATGTTTCCAGAGCTCTCCAGGCCAAC 838

DB 1676 ATAGCCACGGCGATCTCCAGACCAATCTCATGTTTCCAGAGCTCTCCAGGCCAAC 1617

QY 839 CCAATAGCCCTGCTATAGTATGATGACATCTCGGCTTCTAGCTTGTCTCTTAG 898

DB 1616 CCAATAGCCCTGCTATAGTATGATGACATCTCGGCTTCTAGCTTGTCTCTTAG 1557

QY 899 TGTTCCTTAATCAGATTAAGTCTGCTGGAAGCTTTCAATTTTACAGCCCTGAGCAGTCTT 958

DB 1556 TGTTCCTTAATCAGATTAAGTCTGCTGGAAGCTTTCAATTTTACAGCCCTGAGCAGTCTT 1497

QY 959 CTTTGCTAGTGAATTAATGATGATGCTTTTCCGTAATTAAGCAAAATTAAT 1009

DB 1496 CTTTGCTAGTGAATTAATGATGATGCTTTTCCGTAATTAAGCAAAATTAAT 1446

RESULT 10

AAFI5623

ID AAFI5623 standard; cDNA; 1155 BP.

XX

AC AAFI5623;

XX

DT 13-MAR-2001 (first entry)

XX

DE Human prostate cancer antigen nucleotide sequence SEQ ID NO:58.

XX

KM Human; prostate cancer; prostate cancer antigen; detection; diagnosis;

KM neuroprotective; cytostatic; cardiostatic; immunomodulatory; muscular;

KM vulnery; gastrointestinal; nephrotoxic; anti-infective; gynecological;

KM antibacterial; gene therapy; neural; immune; reproductive; renal;

KM gastrointestinal; pulmonary; cardiovascular; proliferative disorder;

KM wound; infectious disease; ss.

XX

OS Homo sapiens.



PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;  
XX WPI; 2003-569235/53.  
DR P-PSDB; ADE07898.  
XX  
PT New polynucleotides, useful for expressing recombinant proteins for  
PT analysis, characterization or therapeutic use, or as markers for tissues  
PT in which the corresponding protein is preferentially expressed.  
XX  
PS Claim 1; SEQ ID NO 53; 1177bp; English.  
XX  
CC The invention comprises the amino acid and coding sequences of novel  
CC proteins. The DNA and protein sequences of the invention are useful as:  
CC markers for tissues in which the corresponding protein is preferentially  
CC expressed; as molecular weight markers on gels; as chromosome markers or  
CC tags; to identify chromosomes or to map related gene positions; and to  
CC compare with endogenous DNA sequences in patients to identify potential  
CC genetic disorders. The present DNA sequence represents a gene of the  
CC invention.  
XX  
SQ Sequence 1560 BP; 444 A; 363 C; 322 G; 431 T; 0 U; 0 Other;  
Query Match 88.2%; Score 905.6; DB 10; Length 1560;  
Best Local Similarity 91.9%; Pred. No. 2e-254;  
Matches 987; Conservative 0; Mismatches 39; Indels 48; Gaps 1;  
QY 2 GGCAGAGTGGGCAAAAAATCAAGTATTTGTCCTCCGGAACAAGCTATCATTAACAG 61  
DB 487 GGCAGATAGAGATCTATTATAGAACTCTTGGCAGTGAACAACCTGTGTCTACAG 546  
QY 62 ATAAACAACCTGATGAGATGTTCCGCCAAGCCCACTATTTTCTTCCTCAATGCTG 121  
DB 547 ATAAACAACCTGATGAGATGTTCCGCCAAGCCCACTATTTTCTTCCTCAATGCTG 606  
QY 122 AAACAAAGCTCCAGAGCTGGAACATACCTTGTCTTCTTGAAGAAATTTTCCCTGATG 181  
DB 607 AAACAAAGCTCCAGAGCTGGAACATACCTTGTCTTCTTGAAGAAATTTTCCCAATA 666  
QY 182 TTATTAAGATACATTTGGCAAGAAAGAACCAACAGATTTCTGGATCCCAAGAGGGA 241  
DB 667 TTATTAAGATACATTTGGCAAGAAAGAACCAACAGATTTCTGGATCCCAAGAGGGA 726  
QY 242 ACAACATGAAGACTAACGACATACATGAATTTAGCTGTTAAGGCTCCAGAAAGT 301  
DB 727 ACAACATGAAGACTAACGACATACATGAATTTAGCTGTTAAGGCTCCAGAAAGT 786  
QY 302 CACTGCAAAAGACACAGATGTATCGTCAGACATGAGAAATTAATAAAGAGAGTATC 361  
DB 787 CACTGCAAAAGACACAGATGTATCGTCAGACATGAGAAATTAATAAAGAGAGTATC 846  
QY 362 AAGAAATTAATCTTCTCCATTAAGC----- 389  
DB 847 AAGAAATTAATCTTCTCCATTAAGC----- 389  
QY 390 -----GGATGTCATCAATGATGCCCAAGCAATTTGTTCAAAAGATG 433  
DB 907 ATTCAAAAGATGCAATGATGTCATCAATGATGCCCAAGCAATTTGTTCAAAAGATG 966  
QY 434 CAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 493  
DB 967 CAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1026  
QY 494 TGCTCTCAAGAGTGTGCTTATTTTGCATCATCCTGCTGTCTGCTTGAAGAACG 553  
DB 1027 TGCTCTCAAGAGTGTGCTTATTTTGCATCATCCTGCTGTCTGCTTGAAGAACG 1086  
QY 554 CTTTCTGCTGAAGAGAGAGAAATCATATAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 613  
DB 1087 CTTTCTGCTGAAGAGAGAGAAATCATATAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1146  
QY 614 TCATCGGTTATTTGCTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 673  
DB 1147 TCATCGGTTATTTGCTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1206

QY 674 TTGGCCATTGAGTTCCTCATGATGATGATGATGATGATGATGATGATGATGATGATG 733  
DB 1207 TTGGCCATTGAGTTCCTCATGATGATGATGATGATGATGATGATGATGATGATGATG 1266  
QY 734 ACCAGTGGGACACAGAGAACTCACTGTGTAATAAAGAGAGAGAGAGAGAGAGAGAG 793  
DB 1267 ACCAGTGGGACACAGAGAACTCACTGTGTAATAAAGAGAGAGAGAGAGAGAGAGAG 1326  
QY 794 TCCAGACCAATCTCTCATGATGATGATGATGATGATGATGATGATGATGATGATG 853  
DB 1327 TCCAGACCAATCTCTCATGATGATGATGATGATGATGATGATGATGATGATGATG 1386  
QY 854 TATAGTGTAGACATCTCGGGGCTTCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 913  
DB 1387 TATAGTGTAGACATCTCGGGGCTTCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1446  
QY 914 TAACTGCTGGAAGCTTTCATTTTACAGGCTTGAAGAGAGAGAGAGAGAGAGAGAG 973  
DB 1447 TAACTGCTGGAAGCTTTCATTTTACAGGCTTGAAGAGAGAGAGAGAGAGAGAGAG 1506  
QY 974 TATGTGTGTGTTTTCCTGATTAAGCAAAATTAATTAATAAAGAGAGAGAGAGAG 1027  
DB 1507 TATGTGTGTGTTTTCCTGATTAAGCAAAATTAATTAATAAAGAGAGAGAGAGAG 1560  
RESULT 12  
ABL5007  
ID ABL5007 standard; DNA; 1421 BP.  
XX  
AC ABL5007;  
XX  
DT 15-MAY-2002 (first entry)  
XX  
DE Lung cancer related gene sequence SEQ ID NO:3344.  
XX  
KW Human: cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytotoxic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;  
KW gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN M0200194629-A2.  
XX  
PD 13-DEC-2001.  
XX  
PF 30-MAY-2001; 2001MO-US010838.  
XX  
PR 05-JUN-2000; 2000US-0209473P.  
PR 05-JUN-2000; 2000US-0209531P.  
PR 18-SEP-2000; 2000US-023133P.  
PR 18-SEP-2000; 2000US-0231617P.  
PR 20-SEP-2000; 2000US-0234009P.  
PR 20-SEP-2000; 2000US-0234034P.  
PR 20-SEP-2000; 2000US-0234052P.  
PR 22-SEP-2000; 2000US-0234509P.  
PR 22-SEP-2000; 2000US-0234567P.  
PR 25-SEP-2000; 2000US-0234923P.  
PR 25-SEP-2000; 2000US-0234924P.  
PR 25-SEP-2000; 2000US-0235077P.  
PR 25-SEP-2000; 2000US-0235082P.  
PR 25-SEP-2000; 2000US-0235134P.  
PR 25-SEP-2000; 2000US-0235280P.  
PR 26-SEP-2000; 2000US-0235637P.  
PR 26-SEP-2000; 2000US-0235638P.  
PR 27-SEP-2000; 2000US-0235711P.  
PR 27-SEP-2000; 2000US-0235720P.  
PR 27-SEP-2000; 2000US-0235840P.  
PR 27-SEP-2000; 2000US-0235863P.  
PR 28-SEP-2000; 2000US-0236028P.  
PR 28-SEP-2000; 2000US-0236032P.  
PR 28-SEP-2000; 2000US-0236033P.  
PR 28-SEP-2000; 2000US-0236033P.



PD 10-SEP-2004.  
 XX  
 PE 22-FEB-2004; 2004WO-DE000433.  
 XX 27-FEB-2003; 2003DE-01009985.  
 PR 14-MAY-2003; 2003DE-01022134.  
 XX  
 PA (HINZ/) HINZMANN B.  
 PA (DAHL/) DAHL E.  
 PA (ROSE/) ROSENTHAL A.  
 PA (HERM/) HERMANN K.  
 PA (PIL/) PILARSKY C.  
 XX  
 PI Hinzmann B, Dahl E, Rosenthal A, Hermann K, Pilarsky C, Specht T;  
 PI Schmitt A, Beckmann G, Bruemendorf T, Kimmernann H, Roeppcke S;  
 PI Kinzhong U, Staub E;  
 XX  
 DR WP1; 2004-653386/63.  
 XX  
 PT New nucleic acids, and encoded proteins, from prostatic cancer tissue,  
 PT useful for diagnosis, treatment and in screening for specific binding  
 PT agents.  
 XX  
 PS Claim 1; Page 282; 1607bp; German.  
 XX  
 CC This invention describes novel cytosolic polynucleotide and polypeptide  
 CC sequences which can be used in a method for diagnosing prostatic cancer  
 CC or the risk of developing prostatic cancer. Diagnosis is based on  
 CC determining over transcription or over expression of the sequences in  
 CC prostatic tissue. Screening for inhibitors of the sequences or detection  
 CC substances involves a binding assay, any compounds that bind are  
 CC selected, optionally after deconvolution of mixtures. Detection of a  
 CC predetermined minimum level of the reporter indicates the presence of  
 CC tumor cells. Inhibitors can be chosen from antisense oligonucleotides,  
 CC short-interfering RNA or ribozymes; an organic molecule of molecular  
 CC weight below 5000, preferably 300, that binds to the polypeptide; an  
 CC aptamer against the polypeptide; a (monoclonal) antibody (Ab) against the  
 CC polypeptide, preferably humanised or human; an anti-idiotypic, non-human  
 CC (monoclonal) antibody directed against Ab or any of the above derivatised  
 CC with a reporter group, cell toxin, immunostimulatory molecules and/or  
 CC radioisotope. The polynucleotides are identified in human prostatic  
 CC cancer by differential expression analysis, using DNA microarrays,  
 CC between normal and tumorous tissues, with (over)expression being detected  
 CC by quantitative PCR. Analysis of prostatic cancer samples showed that  
 CC CD4 was upregulated in many of them. Sections of tissue, isolated from  
 CC prostatic cancer patients, or subjects at risk, were incubated  
 CC biotinylated second antibody; streptavidin-conjugated horseradish  
 CC peroxidase and then diaminobenzidine as colour former (brown). The  
 CC samples were counterstained with hemalum (blue). Malignant cells stained  
 CC strongly but non-malignant cells only weakly. In 15 of 63 samples of  
 CC adenocarcinoma, membrane and cytoplasmic staining was very strong, and  
 CC lymph node metastases were also stained. ADR65805-ADR6654 represent the  
 CC polynucleotide and polypeptide sequences used in the method of the  
 CC invention.  
 XX  
 SQ Sequence 1421 BP; 419 A; 325 C; 292 G; 385 T; 0 U; 0 Other;

Query Match 87.8%; Score 902.2; DB 13; Length 1421;  
 Best Local Similarity 92.8%; Pred. No. 1.9e-253;  
 Matches 977; Conservative 0; Mismatches 28; Indels 48; Gaps 1;

QY 20 AATTCAGGATTTGGTCCCGGAACAAGCTTATCATTAACGATTAACACTTGATGAG 79  
 DB 369 ATAAGAACTCTTTGGCAGTGGAAACAACCTGTGTACAGATTAACAACCTTGATGAG 428  
 QY 80 ATGTTTCCCGCAAGCCCACTATTTTCTTCTCAATGCGAACAAGCTCCAGAAG 139  
 DB 429 ATGTTTCCCGCAAGCCCACTATTTTCTTCTCAATGCGAACAACCTTGATGAG 488  
 QY 140 CTGGAACATACCTTGTCTTCTGAGAAATTTTCCCTGATGTTATTAAGATACATGGC 199  
 DB 489 CTGGAACATACCTTGTCTTCTGAGAAATTTTCCCGATATTAATTAAGATACATGGC 548

QY 200 AAGAAAAGAGAGAAACAGATTCTGGGATCCAGAGAGGGGAAACACCATGAGACTAAG 259  
 DB 549 AAGAAAAGAGAGAAACAGATTCTGGGATCCAGAGAGGGGAAACACCATGAGACTAAG 608  
 QY 260 ACAATATCATGAAATTTAGTGTGTTAACGGTGCAGAAAAGTCACTGAGCAAAAGACACA 319  
 DB 609 ACAATATCATGAAATTTAGTGTGTTAACGGTGCAGAAAAGTCACTGAGCAAAAGACACA 668  
 QY 320 GATGATCGTCAGACATGAGATTAATTAATAACGAGAGTTGATCAAGAAATTAATCTTCTC 379  
 DB 669 GATGATCGTCAGACATGAGATTAATTAATAACGAGAGTTGATCAAGAAATTAATCTTCTC 728  
 QY 380 CAATTAAGAC-----GG 391  
 DB 729 CAATTAAGACAGATGTCACACAGTGGATCCAAAGACAGTTATTCAAAAGATGCAATG 788  
 QY 392 ATGTCATCAATGATGATCCCAAGACAAATGTTCAAAAAGATGCAATGATACACTAGTCG 451  
 DB 789 ATGTCATCAATGATGATCCCAAGACAAATGTTCAAAAAGATGCAATGATACACTAGTCG 848  
 QY 452 TGCAGCTCAAAACACCTCTGCATATTACATGTAACCTCTCTGCTCTCAAGAGTGG 511  
 DB 849 TGCAGCTCAAAACACCTCTGCATATTACATGTAACCTCTCTGCTCTCAAGAGTGG 908  
 QY 512 TCTATTTTGCATCATATGACCTGCTGCTGCTTGAAGAGGGCTTCTGCTGCAATGAG 571  
 DB 909 TCTATTTTGCATCATATGACCTGCTGCTGCTTGAAGAGGGCTTCTGCTGCAATGAG 968  
 QY 572 AGAATCATATAACAGACGATGAGCAACAAGAGGCACTTTTCTCATGCGTTATGTCCT 631  
 DB 969 AGAATCATATAACAGACGATGAGCAACAAGAGGCACTTTTCTCATGCGTTATGTCCT 1028  
 QY 632 AGAAGCTCTTCTGAGGATCTAGTGGGCTTTCTTCTGCGTTTGGCCATTCAGTTCT 691  
 DB 1029 AGAAGCTCTTCTGAGGATCTAGTGGGCTTTCTTCTGCGTTTGGCCATTCAGTTCT 1088  
 QY 692 CATGTGTCTCTTCTCATATTAATGTAATGATTAAGGTTTAAACGATGGGCAACAGAG 751  
 DB 1089 CATGTGTCTCTTCTCATATTAATGTAATGATTAAGGTTTAAACGATGGGCAACAGAG 1148  
 QY 752 AACCTCACTGTAAATTAACATGAGATAGCAAGCGGATCTCCAGACCAATCTCTCC 811  
 DB 1149 AACCTCACTGTAAATTAACATGAGATAGCAAGCGGATCTCCAGACCAATCTCTCC 1208  
 QY 812 ATGTTTCCACAGCTCTCCAGCCCAACCAATAGCGCTGCTAATGATGATGACATCTCG 871  
 DB 1209 ATGTTTCCACAGCTCTCCAGCCCAACCAATAGCGCTGCTAATGATGATGACATCTCG 1268  
 QY 872 CGGCTTACACCTTGTCTCTCTCTTATGATGTTCTTAAATCAATGATGCTGGAAGCTT 931  
 DB 1269 CGGCTTACACCTTGTCTCTCTCTTATGATGTTCTTAAATCAATGATGCTGGAAGCTT 1328  
 QY 932 TCAATTTACAGCGCTGAGACAGCTCTTGTGAGTGAATATGATGATGATGATGATGAT 991  
 DB 1329 TCAATTTACAGCGCTGAGACAGCTCTTGTGAGTGAATATGATGATGATGATGATGAT 1388  
 QY 992 GTAATAGCAAAATTAATTTAAAAAATGAAA 1024  
 DB 1389 GTAATAGCAAAATTAATTTAAAAAATGAAA 1421

RESULT 14  
 ADR66783  
 ID ADR66783 standard; DNA; 1421 BP.  
 XX  
 AC ADR66783;  
 XX  
 DT 02-DEC-2004 (first entry)  
 XX  
 DE Human prostatic carcinoma derived DNA SEQ ID 76 #4.  
 XX  
 KW human; cytostatic; diagnosis; prostatic cancer;

KW differential expression analysis; ds.  
XX Homo sapiens.  
OS MO2004076614-A2.  
XX 10-SEP-2004.  
XX 22-FEB-2004; 2004WO-DE000433.  
XX 27-FEB-2003; 2003DE-0100985.  
PR 14-MAY-2003; 2003DE-01022134.  
XX (HINZ/) HINZMANN B.  
PA (DAHL/) DAHL E.  
PA (ROSE/) ROSENTHAL A.  
PA (HERM/) HERMANN K.  
PA (PILA/) PILARSKY C.  
XX Hinzmann B, Dahl E, Rosenthal A, Hermann K, Pilarsky C, Specht T,  
PI Schmitt A, Beckmann G, Bruemendorf T, Kinneemann H, Roepcke S,  
PI Xinhong L, Staub E,  
XX WPI: 2004-653386/63.  
XX New nucleic acids, and encoded proteins, from prostatic cancer tissue,  
PT useful for diagnosis, treatment and in screening for specific binding  
PT agents.  
XX Claim 1; Page 1452; 1607pp; German.  
XX This invention describes novel cytosolic polynucleotide and polypeptide  
CC sequences which can be used in a method for diagnosing prostatic cancer  
CC or the risk of developing prostatic cancer. Diagnosis is based on  
CC determining over transcription or over expression of the sequences in  
CC prostatic tissue. Screening for inhibitors of the sequences or detection  
CC substances involves a binding assay, any compounds that bind are  
CC selected, optionally after deconvolution of mixtures. Detection of a  
CC predetermined minimum level of the reporter indicates the presence of  
CC tumour cells. Inhibitors can be chosen from antisense oligonucleotides,  
CC short-interfering RNA or ribozymes; an organic molecule of molecular  
CC weight below 5000, preferably 300, that binds to the polypeptide; an  
CC aptamer against the polypeptide; a (monoclonal) antibody (Ab) against the  
CC polypeptide, preferably humanised or human; an anti-idiotype, non-human  
CC (monoclonal) antibody directed against Ab or any of the above derivatised  
CC with a reporter group, cell toxin, immunostimulatory molecules and/or  
CC radioisotope. The polynucleotides are identified in human prostatic  
CC cancer by differential expression analysis, using DNA microarrays,  
CC between normal and tumorous tissues, with (over)expression being detected  
CC by quantitative PCR. Analysis of prostatic cancer samples showed that  
CC CD24 was upregulated in many of them. Sections of tissue, isolated from  
CC prostatic cancer patients, or subjects at risk, were incubated  
CC sequentially with anti-human CD4 murine monoclonal antibodies;  
CC biotinylated second antibody; streptavidin-conjugated horseradish  
CC peroxidase and then diaminobenzidine as colour former (brown). The  
CC samples were counterstained with hemalum (blue). Malignant cells stained  
CC strongly but non-malignant cells only weakly. In 15 of 63 samples of  
CC adenocarcinoma, membranes and cytoplasmic staining was very strong, and  
CC lymph node metastases were also stained. ADR65805-ADR65954 represent the  
CC polynucleotide and polypeptide sequences used in the method of the  
CC invention.  
XX  
XX Sequence 1421 BP; 419 A; 325 C; 292 G; 385 T; 0 U; 0 Other;  
SQ  
Query Match 87.8%; Score 902.2; DB 13; Length 1421;  
Best Local Similarity 92.8%; Pred. No. 1.9e-253;  
Matches 977; Conservative 0; Mismatches 28; Indels 48; Gaps 1;  
QY 20 AATCAAGGTAATTTGGTCCCGAACAAGCTTATCATTAACAAGCACTTGATGCG 79  
DB 369 ATTAAGAACTCTTTGGCAGTGGAAACAACACTTGTGTCAAGATTAACAACCTTATGACG 428  
QY 80 ATGTTTCCCAAGCCCACTATTTTCTTCTCAATTGCTGAACAAAGCTCCAGAAAG 139

DB 429 ATGTTTCCCAAGCCCACTATTTTCTTCTCAATTGCTGAACAAAGCTCCAGAAAG 488  
QY 140 CTGGAACATACCTTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAGATACATTTGGC 199  
DB 489 CTGGAACATACCTTTGTCTTCTTGAGAAATTTTCCAGATTAATTAATTAAGATACATTTGGC 548  
QY 200 AAGAAAAGAAAGCAACAAGATTTGGGATCCAGAGAGGGGAAACACATGAAGACTAACG 259  
DB 549 AAGAAAAGAAAGCAACAAGATTTGGGATCCAGAGAGGGGAAACACCATGAAGACTAACG 608  
QY 260 ACACATGATGAAGAAATTTAGCTGTGTTAACGGTGCAGAAAGTCACTGGAACAAAGAACCA 319  
DB 609 ACACATGATGAAGAAATTTAGCTGTGTTAACGGTGCAGAAAGTCACTGGAACAAAGAACCA 668  
QY 320 GATGATGCTCAGACATGAGAAATTAATTAAGAGAGGATGATCAAGAAATTAATTTCTCTC 379  
DB 669 GATGATGCTCAGACATGAGAAATTAATTAAGAGAGGATGATCAAGAAATTAATTTCTCTC 728  
QY 380 CAATTAAGAC-----GG 391  
DB 729 CAATTAAGACAGATGTCACCAAGTGGATCCAAAGACAGTTATTCAAAAGATGCAATG 788  
QY 392 ATGTCATCAAAATGGAATCCAAAGACAAATTTGTTCAAAAGATGCAAAATGATACACTATGC 451  
DB 789 ATGTCATCAAAATGGAATCCAAAGACAAATTTGTTCAAAAGATGCAAAATGATACACTATGC 848  
QY 452 TGCACTCACAACACACCTCTGCAATTAATTAACATGTAACCTCTCTGCTCTCAAGAGTGG 511  
DB 849 TGCACTCACAACACACCTCTGCAATTAATTAACATGTAACCTCTCTGCTCTCAAGAGTGG 908  
QY 512 TCTATTTTGCATCATCATCTGCTGCTGCTTGAAGACAGGCTTTGCTGCAATGAG 571  
DB 909 TCTATTTTGCATCATCATCTGCTGCTGCTTGAAGACAGGCTTTGCTGCAATGAG 968  
QY 572 AGAAATATAACAGCGGTGGCAAGAGGCACTTTTCTCATGAGTTATGTCCT 631  
DB 969 AGAAATATAACAGCGGTGGCAAGAGGCACTTTTCTCATGAGTTATGTCCT 1028  
QY 632 AGAAGCTCTTTCGAGAGATCATGTTGGGCTTTCTTCTGGGTTGGGCCATTTCAATTCT 691  
DB 1029 AGAAGCTCTTTCGAGAGATCATGTTGGGCTTTCTTCTGGGTTGGGCCATTTCAATTCT 1088  
QY 692 CATGTGTGTAATCTATTCATTAATTTGATTAACGTTTTCACACAGTGGACACAGAG 751  
DB 1089 CATGTGTGTAATCTATTCATTAATTTGATTAATGTTTTCACACAGTGGACACAGAG 1148  
QY 752 AACCTCACTGTAATTAACAATGAGAAAGCCAGCGGAGATTCACACCAATCTCTCC 811  
DB 1149 AACCTCACTGTAATTAACAATGAGAAAGCCAGCGGAGATTCACACCAATCTCTCC 1208  
QY 812 ATGTTTTCACAGCTCTCCAGCCCAACCAATAGCCCTGCTTAATGTAAGATCATCTCG 871  
DB 1209 ATGTTTTCACAGCTCTCCAGCCCAACCAATAGCCCTGCTTAATGTAAGATCATCTCG 1268  
QY 872 CGGCTTTCAGCTTGTCCCTCTTCTTGAAGTTCTTTAATCAAGATTAAGCTGGAAGCTT 931  
DB 1269 CGGCTTTCAGCTTGTCCCTCTTCTTGAAGTTCTTTAATCAAGATTAAGCTGGAAGCTT 1328  
QY 932 TCAATTTTACAGCGCTTGAAGAGCTTCTTGTGTAATTAATTAATGAGTGTGTTTCC 991  
DB 1329 TCAATTTTACAGCGCTTGAAGAGCTTCTTGTGTAATTAATTAATGAGTGTGTTTCC 1388  
QY 992 GTAATTAAGCAAAATTAATTTTAAATAATGAATA 1024  
DB 1389 GTAATTAAGCAAAATTAATTTTAAATAATGAATA 1421  
RESULT 15  
ABK64529  
ID ABK64529 standard; DNA; 1586 BP.  
XX  
AC ABK64529;

XX 18-JUN-2002 (first entry)  
 DT Human benign prostatic hyperplasia gene #424.  
 XX Human benign prostatic hyperplasia; BPH; prostate cancer; gene; ds.  
 XX Homo sapiens.  
 OS MO200212440-A2.  
 XX 14-FEB-2002.  
 PD 07-AUG-2001; 2001WO-US024708.  
 XX 07-AUG-2001; 2000US-0223323P.  
 XX 05-JUN-2001; 2001US-00873319.  
 PR (GENE-) GENE LOGIC INC.  
 XX (NISR) JAPAN TOBACCO INC.  
 PA Munger WE, Kulkarni P, Getzenberg RH, Waga I, Yamamoto J;  
 PI WPI; 2002-257476/30.  
 DR Identifying drugs for and diagnosing benign prostatic hyperplasia, by  
 XX detecting expression levels of one or more genes in prostate cells from  
 PT patient that are differentially regulated compared to normal prostate  
 PT cells.  
 XX Disclosure; Page 247-248; 444pp; English.  
 PS The invention relates to a method of diagnosing (I) the onset or  
 CC progression of benign prostatic hyperplasia (BPH), or screening (II) for  
 CC or identifying an agent that modulates the onset or progression of BPH.  
 CC The method is based on changes in gene expression in BPH tissue isolated  
 CC from patients exhibiting different clinical states of prostate  
 CC hyperplasia as compared to normal prostate tissue. (I) comprises  
 CC detecting the expression levels of one or more genes in prostate cells  
 CC from the subject that are differentially regulated compared to normal  
 CC prostate cells. (II) comprises preparing a first gene expression profile  
 CC of BPH cells or BPH-like cell population, exposing the cells to the  
 CC agent, preparing a second gene expression profile of the agent exposed  
 CC cells, and comparing the first and second gene expression profiles. (I)  
 CC is useful for diagnosing the onset or progression of BPH. (II) is useful  
 CC for identifying an agent that modulates the onset or progression of BPH.  
 CC The methods are useful to present information identifying the expression  
 CC level in a tissue or cells, by comparing the expression level of genes  
 CC given in the specification in the tissue or cells to the level of  
 CC expression of gene in the database, and displaying the expression levels  
 CC of at least one gene in the tissue or cell sample compared to the  
 CC expression level in BPH. Agents using (II) are useful for treating BPH or  
 CC prostate cancer. ABK64106-ABK64860 represent human benign prostatic  
 CC hyperplasia gene sequences of the invention  
 CC  
 XX  
 SQ Sequence 1586 BP; 449 A; 369 C; 330 G; 438 T; 0 U; 0 Other;  
 Query Match 81.6%; Score 838; DB 6; Length 1586;  
 Best Local Similarity 88.4%; Pred. No. 1.3e-234;  
 Matches 962; Conservative 0; Mismatches 30; Indels 96; Gaps 1;

QY 200 AAGAAAAGAGACACGATTCGGGATCCAGAGGGGAAACCATGAGACTAAG 259  
 DB 679 AAGAAAAGAGAGACACGATTCGGGATCCAGAGGGGAAACCATGAGACTAAG 738  
 QY 260 ACACATCATGAAATTTAGCTGTGTTACGGTCCAGAAAAGTCACTGACAAAGACACA 319  
 DB 739 ACACATCATGAAATTTAGCTGTGTTACGGTCCAGAAAAGTCACTGACAAAGACACA 798  
 QY 320 GATGTATCGTCAGCATGAGAAATTAATAAAACGAGGTGATCAAGAAATTAATCTTCTC 379  
 DB 799 GATGTATCGTCAGCATGAGAAATTAATAAAACGAGGTGATCAAGAAATTAATCTTCTC 858  
 QY 380 CAATTAAGAC-----GATGTATCAAA 389  
 DB 859 CAATTAAGACAGATGTACACACAGTGTATCCAAAGACAGTTATTAAGATCAATG 918  
 QY 390 -----GATGTATCAAA 403  
 DB 919 ATGTACACAGATGTATCCAAATTAATAAAAGATGCAAAATGATGTATCAACAA 978  
 QY 404 TGGATCCCAAGACAAATTTGTTCAAAAAGATGCAAAATGATGTATCAACAA 463  
 DB 979 TGGATCCCAAGACAAATTTGTTCAAAAAGATGCAAAATGATGTATCAACAA 1038  
 QY 464 ACACTCTGCATATTACATGATCTCTCTGCTCTCAAGAGTGTGATATTGCA 523  
 DB 1039 ACACTCTGCATATTACATGATCTCTCTGCTCTCAAGAGTGTGATATTGCA 1098  
 QY 524 TCATCAGCTGCTGTCTGTGAAAGAGCGTTTCTGCTGATGAGAAATCATPAC 583  
 DB 1099 TCATCAGCTGCTGTCTGTGAAAGAGCGTTTCTGCTGATGAGAAATCATPAC 1158  
 QY 584 AGACGGGGGACAAAGAGGCGATCTTTCGATCGGTTATTTGCTCAGAGGCTCTC 643  
 DB 1159 AGACGGGGGACAAAGAGGCGATCTTTCGATCGGTTATTTGCTCAGAGGCTCTC 1218  
 QY 644 TGAAGATCTAGTGGGCTTTCTTCTGGGTTGGGCCATTTCAGTGTGTAAT 703  
 DB 1219 TGAAGATCTAGTGGGCTTTCTTCTGGGTTGGGCCATTTCAGTGTGTAAT 1278  
 QY 704 ATTCTATCATATTATGTTAAACGTTTCAAAACAGTGGGACACAGAAACCTCATCTG 763  
 DB 1279 ATTCTATCATATTATGTTAAACGTTTCAAAACAGTGGGACACAGAAACCTCATCTG 1338  
 QY 764 TAATTAACAATGAGAAATAGCCAGCGATCTCCAGACCAATCTCTCAATCTTTCCACA 823  
 DB 1339 TAATTAACAATGAGAAATAGCCAGCGATCTCCAGACCAATCTCTCAATCTTTCCACA 1398  
 QY 824 GCTCCTCAGCAACCCAAATAGGCGCTGCTATGATGATGATGATGATGATGATGATGAT 883  
 DB 1399 GCTCCTCAGCAACCCAAATAGGCGCTGCTATGATGATGATGATGATGATGATGATGAT 1458  
 QY 884 TTGTCCCTCTTCTAGTGTCTTTAATGATGATGATGATGATGATGATGATGATGATGAT 943  
 DB 1459 TTGTCCCTCTTCTAGTGTCTTTAATGATGATGATGATGATGATGATGATGATGATGAT 1518  
 QY 944 CCTGGAAGAGCTCTTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1003  
 DB 1519 CCTGGAAGAGCTCTTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1578  
 QY 1004 ATAAATTT 1011  
 DB 1579 ATAAATTT 1586

Search completed: December 10, 2005, 20:46:12  
 Job time : 671 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OW nucleic - nucleic search, using bw model

Run on: December 10, 2005, 19:57:47 ; Search time 5312 Seconds  
(without alignments)  
10989.877 Million cell updates/sec

Title: US-10-031-158B-13

Perfect score: 1027  
Sequence: 1 gggcagaagctggcgcaaaaa.....attcaaaaaatgaaaaagct 1027

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

GenEmbl: \*  
1: gb ba: \*  
2: gb in: \*  
3: gb env: \*  
4: gb om: \*  
5: gb ov: \*  
6: gb pat: \*  
7: gb ph: \*  
8: gb pr: \*  
9: gb ro: \*  
10: gb sts: \*  
11: gb sy: \*  
12: gb un: \*  
13: gb vi: \*  
14: gb ntg: \*  
15: gb pl: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1027	100.0	1027	6	CQ983344 Sequence
2	1027	100.0	1027	8	AF151103 Homo sapi
3	1025.4	99.8	1027	6	AX074415 Sequence
4	1010	98.3	1825	8	BC072396 Homo sapi
5	1008.8	98.2	1054	8	BC062761 Homo sapi
6	984	95.8	1530	6	CS032596 Sequence
7	984	95.8	1530	6	CS041548 Sequence
8	984	95.8	1530	6	HUMTCRGA
9	973	94.7	1470	8	BC072387 Homo sapi
10	938	91.3	958	8	HUMTCRGAD
11	903.8	88.0	1655	8	BC039725 Homo sapien
12	902.2	87.8	1421	6	CO896252 Sequence
13	902.2	87.8	1421	6	AX332835 Sequence
14	902.2	87.8	1421	8	HUMTCGAX
15	877.4	85.4	1402	8	HUMTCRGAA
16	865.2	84.2	1402	8	HUMTCRGA
17	857.4	83.5	1003	8	HUMTCRGAAB
18	842.4	82.0	1757	8	BC039116 Homo sapi

19	838	81.6	1586	6	CS032845 Sequence
20	838	81.6	1586	6	CS041797 Sequence
21	838	81.6	1586	6	HUMTCRGAD
22	821.6	80.0	1046	8	HUMTCGCH
23	821.6	80.0	1046	10	G31759 sKWS1280 Et
24	795.8	77.5	1162	6	CQ489650 Sequence
25	795.8	77.5	1162	6	CQ492275 Sequence
26	795.8	77.5	1162	6	CQ492662 Sequence
27	795.8	77.5	1162	6	CQ492666 Sequence
28	795.8	77.5	1162	6	CQ492973 Sequence
29	795.8	77.5	1162	6	CQ493284 Sequence
30	795.8	77.5	1162	6	CQ495492 Sequence
31	795.8	77.5	1162	6	CQ495873 Sequence
32	733.4	71.4	927	8	HUMTCGCT
33	619	60.3	924	6	CQ493214 Sequence
34	601	58.5	786	8	HUMTCRGAAC
35	589	57.4	847	8	HUMTCGCD3
36	589	57.4	140691	8	AF159056
37	589	57.4	171816	8	AC006033
38	579.6	56.4	586	8	HUMTCGAX
39	578.2	56.3	1041	8	AY190025
40	577.2	56.2	1041	8	AY190027
41	576.4	56.1	6658	14	AC130306
42	522.8	50.9	1080	6	CS036408
43	522.8	50.9	1080	6	CS045360
44	522.8	50.9	1080	8	HSTCRGR
45	516.4	50.3	1080	6	AR364341 Sequence

#### ALIGNMENTS

RESULT 1  
LOCUS CQ983344 1027 bp DNA linear PAT 25-JAN-2005  
DEFINITION Sequence 2 from Patent WO2005000889.  
ACCESSION CQ983344  
VERSION CQ983344.1 GI:58191770

KEYWORDS  
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.

REFERENCE  
AUTHORS Berzofsky, J.A., Oh, S. and Pastan, I.  
TITLE Immunogenic peptides for the treatment of prostate and breast cancer

JOURNAL Patent: WO 2005000889-A 2 06-JAN-2005;  
National Institutes of Health (US)  
Location/Qualifiers

FEATURES  
source

1..1027  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

#### ORIGIN

Query Match 100.0%; Score 1027; DB 6; Length 1027;  
Best Local Similarity 100.0%; Pred. No. 3.5e-253;  
Matches 1027; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GGGCAAGCTGGGCAAAAATCAAGCTTTTGTCCCGGAACAAAGCTTATCATACA	60
DB	1	GGGCAAGCTGGGCAAAAATCAAGCTTTTGTCCCGGAACAAAGCTTATCATACA	60
QY	61	GATTAACAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	120
DB	61	GATTAACAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	120
QY	121	GAACAAGCTGCAAGAGCTGCAAGAGCTGCAAGAGCTGCAAGAGCTGCAAGAGCT	180
DB	121	GAACAAGCTGCAAGAGCTGCAAGAGCTGCAAGAGCTGCAAGAGCTGCAAGAGCT	180

```

OY 181 GTTATTGAATACATTGGCAGAAAGAAAGCAACAGATTCTGGGATCCGAGAGGG 240
Db 181 GTTATTGAATACATTGGCAGAAAGAAAGCAACAGATTCTGGGATCCGAGAGGG 240
OY 241 AACCCATGAAGACTTAAGCAACATACATGAAATTTAGCTGGTTAAGGTCGCAAGAAAG 300
Db 241 AACCCATGAAGACTTAAGCAACATACATGAAATTTAGCTGGTTAAGGTCGCAAGAAAG 300
OY 301 TCACCTGCAAAAGAACACAGATGTATCCCTGACACATGAGAAATTAATTAAGGAGTTGAT 360
Db 301 TCACCTGCAAAAGAACACAGATGTATCCCTGACACATGAGAAATTAATTAAGGAGTTGAT 360
OY 361 CAAGAAATTAATCTTCTCCCAATTAAGACGATGTATCAATGATGATCCCAAGACAT 420
Db 361 CAAGAAATTAATCTTCTCCCAATTAAGACGATGTATCAATGATGATCCCAAGACAT 420
OY 421 TGTTCAAAAGATGCAAAATGATACACTAGCTGACGCTCAAAACACTCTGCATATTAC 480
Db 421 TGTTCAAAAGATGCAAAATGATACACTAGCTGACGCTCAAAACACTCTGCATATTAC 480
OY 481 ATGTACCTCTCTGCTGCTCTCAAGAGTGTGTATTTTGGCATCATCAGCTGTGTCTG 540
Db 481 ATGTACCTCTCTGCTGCTCTCAAGAGTGTGTATTTTGGCATCATCAGCTGTGTCTG 540
OY 541 CTTAGAGACAGCGCTTTCTGCTGCAATGAGAGAAATCATTAACAGCGTGGCAACAGGA 600
Db 541 CTTAGAGAGACAGCGCTTTCTGCTGCAATGAGAGAAATCATTAACAGCGTGGCAACAGGA 600
OY 601 GGGCATCTTTCTTCATGCTGTTATTTGCTTGAAGAGCTTTCTGAGAGATCTAGTTGGGC 660
Db 601 GGGCATCTTTCTTCATGCTGTTATTTGCTTGAAGAGCTTTCTGAGAGATCTAGTTGGGC 660
OY 661 TTTCTTCTGGGTTTGGGCAATTCATGTTCTCATGTTGTTACTATTTATCATTAATTGTA 720
Db 661 TTTCTTCTGGGTTTGGGCAATTCATGTTCTCATGTTGTTACTATTTATCATTAATTGTA 720
OY 721 TAACGCTTTTCAACCAAGTGGGACACAGAGAACTCTCTGTATTAACAATGAGGAT 780
Db 721 TAACGCTTTTCAACCAAGTGGGACACAGAGAACTCTCTGTATTAACAATGAGGAT 780
OY 781 AGCCAGCGGAGATCTCCAGACCAATCTCCATGTTTCCACAGCTCTCCAGCCAAACC 840
Db 781 AGCCAGCGGAGATCTCCAGACCAATCTCCATGTTTCCACAGCTCTCCAGCCAAACC 840
OY 841 AAATAGCCCTGCTATGATGATGATCATCTCGGCTTGAAGCTTCTCTTGAATG 900
Db 841 AAATAGCCCTGCTATGATGATGATCATCTCGGCTTGAAGCTTCTCTTGAATG 900
OY 901 TTTCTTAAATCAGATTAAGCTGGAAGCTTCAATTTTAAACAGCCCTGAAGAGTCTTCT 960
Db 901 TTTCTTAAATCAGATTAAGCTGGAAGCTTCAATTTTAAACAGCCCTGAAGAGTCTTCT 960
OY 961 TTGCTAGTTGAATTAATGAGTGTGTTTTCCTAATAAGCAAAATTAATTAAGGATG 1020
Db 961 TTGCTAGTTGAATTAATGAGTGTGTTTTCCTAATAAGCAAAATTAATTAAGGATG 1020
OY 1021 AAAAGTT 1027
Db 1021 AAAAGTT 1027

```

```

RESULT 2
LOCUS AF151103 1027 bp mRNA linear PRI 03-NOV-2000
DEFINITION Homo sapiens TCRgamma alternate reading frame protein (TCRG) mRNA,
ACCESSION AF151103
VERSION AF151103.1 GI:5758136
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

```

```

REFERENCE
AUTHORS Hominiidae; Homo.
1 (sites)
Bessand,M., Vasmatazis,G., Brinkmann,U., Dury,P., Lee,B. and
Pastan,I.
TITLE High expression of a specific T-cell receptor gamma transcript in
epithelial cells of the prostate
Proc. Natl. Acad. Sci. U.S.A. 96 (16), 9287-9292 (1999)
JOURNAL
PUBMED 10430935
REFERENCE
AUTHORS 2 (bases 1 to 1027)
Wolfgang,C.D., Bessand,M., Vincent,J.J., Lee,B. and Pastan,I.
TITLE TRAP: a nuclear protein expressed in prostate and breast cancer
cells derived from an alternate reading frame of the T cell
receptor gamma chain locus
Proc. Natl. Acad. Sci. U.S.A. 97 (17), 9437-9442 (2000)
JOURNAL
PUBMED 10931945
REFERENCE
AUTHORS 3 (bases 1 to 1027)
Bessand,M., Vasmatazis,G., Brinkmann,U., Dury,P., Lee,B. and
Pastan,I.
TITLE Direct Submission
JOURNAL Submitted (13-MAY-1999) Division of Basic Sciences, Laboratory of
Molecular Biology, National Institutes of Health, National Cancer
Institute, 37 Convent Drive, Bethesda, MD 20892, USA
FEATURES
source
location/Qualifiers
1..1027
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/note="transcript from an unarranged TCRg locus"
1..1027
/gene="TCRG"
1..60
/gene="TCRG"
/notes="gamma 1.2"
61..390
/gene="TCRG"
/notes="gamma 1"
corresponds to exon I"
74..250
/gene="TCRG"
/notes="TARP"
/codon_start=1
/product="TCRgamma alternate reading frame protein"
/protein_id="AA029337.1"
/db_xref="GI:11093514"
/translation="MQWPPSPLEFFQLKQSSRLRLEHFFVLFNFSMLRLYIGKK
RRARFMDPRGRP"
391..438
/gene="TCRG"
/notes="gamma 1; corresponds to exon II"
439..579
/gene="TCRG"
/notes="gamma 1; corresponds to exon III"
ORIGIN
Query Match 100.0%; Score 1027; DB 8; Length 1027;
Best Local Similarity 100.0%; Pred. No. 3.5e-253; Indels 0; Gaps 0;
Matches 1027; Conservative 0; Mismatches 0;
C_region
C_region
/gene="TCRG"
/notes="gamma 1; corresponds to exon II"
439..579
/gene="TCRG"
/notes="gamma 1; corresponds to exon III"
OY 1 GGGCAAGAGTTGGGCAAAATCAAGTATTTGTCGCGGACAAAGCTTATCATTTCA 60
Db 1 GGGCAAGAGTTGGGCAAAATCAAGTATTTGTCGCGGACAAAGCTTATCATTTCA 60
OY 61 GATAAACAATTGATGAGATGTTTCCCAAGCCCACTATTTTCTTCAATTGCT 120
Db 61 GATAAACAATTGATGAGATGTTTCCCAAGCCCACTATTTTCTTCAATTGCT 120
OY 121 GAAACAAAGCTCCGAAAGCTGGAACATACCTTTGTCTTCTTGAAGAAATTTTCCCTGAT 180
Db 121 GAAACAAAGCTCCGAAAGCTGGAACATACCTTTGTCTTCTTGAAGAAATTTTCCCTGAT 180
OY 181 GTTATTGAATACATTGGCAGAAAGAAAGCAACAGATTCTGGGATCCGAGAGGG 240
Db 181 GTTATTGAATACATTGGCAGAAAGAAAGCAACAGATTCTGGGATCCGAGAGGG 240

```

QY 241 AACACATGAAGACTAAGACACATACATGAATTTAGTGTAAACGGTCCAGAAAAG 300  
 DB 241 AACACATGAAGACTAAGACACATACATGAATTTAGTGTAAACGGTCCAGAAAAG 300  
 QY 301 TCACCTGACAAAGAACACAGATGATCTGACACATGAGATATATAAAACGAGTTGAT 360  
 DB 301 TCACCTGACAAAGAACACAGATGATCTGACACATGAGATATATAAAACGAGTTGAT 360  
 QY 361 CAGAAATATATCTTTCTCTCCATTAAGACGGATGTCTACATGATCTCCAAAGACAT 420  
 DB 361 CAGAAATATATCTTTCTCTCCATTAAGACGGATGTCTACATGATCTCCAAAGACAT 420  
 QY 421 TGTTCAAAAGATGCAATGATATACCTACTGCTGACGCTCAACACCTCTGCATATTAC 480  
 DB 421 TGTTCAAAAGATGCAATGATATACCTACTGCTGACGCTCAACACCTCTGCATATTAC 480  
 QY 481 ATGTAACCTCTCTCTCTCTCTCAAGATGTGTATTTTTCATATTCATCATCTGCTGTG 540  
 DB 481 ATGTAACCTCTCTCTCTCTCTCAAGATGTGTATTTTTCATATTCATCATCTGCTGTG 540  
 QY 541 CTTTGAAGAAACGGCTTTCTGCTGCAATGAGAGAAATCATTAACAGCGTGGCAAGGA 600  
 DB 541 CTTTGAAGAAACGGCTTTCTGCTGCAATGAGAGAAATCATTAACAGCGTGGCAAGGA 600  
 QY 601 GGCCATCTTTTCTCATCGGTTATGTCCCTAGAAAGGCTCTTCGAGGATCTAGTTGGGC 660  
 DB 601 GGCCATCTTTTCTCATCGGTTATGTCCCTAGAAAGGCTCTTCGAGGATCTAGTTGGGC 660  
 QY 661 TTTCTTTCTGGGTTTGGGCAATTCAGTTCTCATGTGTACTATTTCTATCATTTATTTGA 720  
 DB 661 TTTCTTTCTGGGTTTGGGCAATTCAGTTCTCATGTGTACTATTTCTATCATTTATTTGA 720  
 QY 721 TTAACGTTTTCACCAACGATGGGCAACAGAGAACCTCACTCTGTATTAACAATGAGGAAT 780  
 DB 721 TTAACGTTTTCACCAACGATGGGCAACAGAGAACCTCACTCTGTATTAACAATGAGGAAT 780  
 QY 781 AGCCAGCGGATCTCCAGACCAATCTCTCATGTGTTCACAGCTCTCCAGCAACCC 840  
 DB 781 AGCCAGCGGATCTCCAGACCAATCTCTCATGTGTTCACAGCTCTCCAGCAACCC 840  
 QY 841 AAATAGCGCTGCTATAGTATGATGACATCTGCGGCTTCAGCCTTGTCCCTCTTTAGTG 900  
 DB 841 AAATAGCGCTGCTATAGTATGATGACATCTGCGGCTTCAGCCTTGTCCCTCTTTAGTG 900  
 QY 901 TTTCTTAATCAGATTAATCTGCTGAGAGCTTTTCAATTTTAAACGCGCTGAAGACGTTCT 960  
 DB 901 TTTCTTAATCAGATTAATCTGCTGAGAGCTTTTCAATTTTAAACGCGCTGAAGACGTTCT 960  
 QY 961 TTGCTATTGAATTAATGTGTGTGTCTTTCCGTAATAAGCAAAATTAATTAATAAATG 1020  
 DB 961 TTGCTATTGAATTAATGTGTGTGTCTTTCCGTAATAAGCAAAATTAATTAATAAATG 1020  
 QY 1021 AAAAGTT 1027  
 DB 1021 AAAAGTT 1027  
 RESULT 3  
 LOCUS AX074415 1027 bp DNA linear PAT 06-FEB-2001  
 DEFINITION Sequence 13 from Patent WO0104309.  
 ACCESSION AX074415  
 VERSION AX074415.1 GI:12710553  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.  
 REFERENCE  
 AUTHORS Patai, I., Bessand, M., Lee, B., Vaarmatzis, G. and Wolfgang, C.  
 TITLE T-cell receptor  $\gamma$ (g) alternate reading frame protein, (tarp) and  
 uses thereof

JOURNAL Patent: WO 0104309-A 13 18-JAN-2001;  
 UNITED STATES GOVERNMENT (US)  
 FEATURES Location/Qualifiers  
 source 1..1027  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 74..250  
 CDS  
 /note="unnamed protein product; Coding region for PS-TCR  
 gamma 1 polypeptide (TARF)"  
 /codon\_start=1  
 /protein\_id="CAC28473.1"  
 /db\_xref="GI:12710554"  
 /translation="MQMFPSPPLFFPLQLKQSRRLKHTFVFLNFSMLRLKIGKK  
 RRAIRFMDRRGRTP"  
 247..582  
 /note="unnamed protein product; Coding region for PS-TCR  
 gamma 2 polypeptide (deduced amino acid sequence not  
 displayed along with DNA sequence, due to overlapping  
 CDS's)"  
 /codon\_start=1  
 /protein\_id="CAC28474.1"  
 /db\_xref="GI:12710555"  
 /translation="MKTNDTYMKFSLTVPEKSLDKHRICIVRHNNKNGVDQELIFP  
 PIKTDIVITDPPDNCSKIDANDTLILQLTWTISAYMYLILLLKSVYFAIITCCLLRT  
 AFCNGKNS"  
 ORIGIN  
 Query Match 99.8%; Score 1025.4; DB 6; Length 1027;  
 Best Local Similarity 99.9%; Pred. No. 9.1e-253;  
 Matches 1026; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GGGCAAGAGTTGGGCAAAAAAATCAAGATTTTGTGCTCCCGGAACAAAGCTTATCATTTACA 60  
 DB 1 GGGCAAGAGTTGGGCAAAAAAATCAAGATTTTGTGCTCCCGGAACAAAGCTTATCATTTACA 60  
 QY 61 GATTAACAACCTTGATGATGATGTTTCCCAAGCCCACTATTTTCTTCCTTCAATTTGCT 120  
 DB 61 GATTAACAACCTTGATGATGATGTTTCCCAAGCCCACTATTTTCTTCCTTCAATTTGCT 120  
 QY 121 GAAACAAGCTCCAGAAAGCTGGACATACCTTTGTTCTTGAAGAAATTTTCCCTGAT 180  
 DB 121 GAAACAAGCTCCAGAAAGCTGGACATACCTTTGTTCTTGAAGAAATTTTCCCTGAT 180  
 QY 181 GTTATTAAGATACATTTGGGAAAGAAAGAGCAACGATTTCTGGATCCCGAGAGGGG 240  
 DB 181 GTTATTAAGATACATTTGGGAAAGAAAGAGCAACGATTTCTGGATCCCGAGAGGGG 240  
 QY 241 AACACATGAAGACTAAGACACATACATGAATTTAGCTGTTAACGGTCCAGAAAAG 300  
 DB 241 AACACATGAAGACTAAGACACATACATGAATTTAGCTGTTAACGGTCCAGAAAAG 300  
 QY 301 TCACCTGACAAAGAACACAGATGATCTGACACATGAGATATATAAAACGAGTTGAT 360  
 DB 301 TCACCTGACAAAGAACACAGATGATCTGACACATGAGATATATAAAACGAGTTGAT 360  
 QY 361 CAGAAATATATCTTTCTCTCCATTAAGACGGATGTCTACATGATCTCCAAAGACAT 420  
 DB 361 CAGAAATATATCTTTCTCTCCATTAAGACGGATGTCTACATGATCTCCAAAGACAT 420  
 QY 421 TGTTCAAAAGATGCAATGATATACCTACTGCTGACGCTCAACACCTCTGCATATTAC 480  
 DB 421 TGTTCAAAAGATGCAATGATATACCTACTGCTGACGCTCAACACCTCTGCATATTAC 480  
 QY 481 ATGTAACCTCTCTCTCTCTCTCAAGATGTGTATTTTTCATATTCATCATCTGCTGTG 540  
 DB 481 ATGTAACCTCTCTCTCTCTCTCAAGATGTGTATTTTTCATATTCATCATCTGCTGTG 540  
 QY 541 CTTTGAAGAAACGGCTTTCTGCTGCAATGAGAGAAATCATTAACAGCGTGGCAAGGA 600  
 DB 541 CTTTGAAGAAACGGCTTTCTGCTGCAATGAGAGAAATCATTAACAGCGTGGCAAGGA 600  
 QY 601 GGCCATCTTTTCTCATCGGTTATGTCCCTAGAAAGGCTTCTGAGGATCTAGTTGGGC 660

```

Db      601 GGCACATCTTTCCTATGCGTATGTCCTTGAAGAGCCTTCTGAGATCTAGTGGGC 660
Oy      661 TTTCTTCTGGGTTTGGGCAATTTGAGTTCATGTGTACTATTTATCATTTATGTA 720
Db      661 TTTCTTCTGGGTTTGGGCAATTTGAGTTCATGTGTACTATTTATCATTTATGTA 720
Oy      721 TAAAGGTTTCAACCAAGTGGGACACAGAGAACTCCTGTATATACATGAGAAAT 780
Db      721 TAAAGGTTTCAACCAAGTGGGACACAGAGAACTCCTGTATATACATGAGAAAT 780
Oy      781 AGCCAGCGGAGATCTCCAGACCAATCTCCATGTTTCCACAGTCTCTCCAGCAACC 840
Db      781 AGCCAGCGGAGATCTCCAGACCAATCTCCATGTTTCCACAGTCTCTCCAGCAACC 840
Oy      841 AAATAGCCGCTGCTATGATGTAGACATCTGGGCTTCTAGCCTTCTCTTCTAGTG 900
Db      841 AAATAGCCGCTGCTATGATGTAGACATCTGGGCTTCTAGCCTTCTCTTCTAGTG 900
Oy      901 TTTCTTAAATCAGATTAAGTGCCTGGAAGCCTTTCATTTTACAGCCCTGAGAGAGTCTCT 960
Db      901 TTTCTTAAATCAGATTAAGTGCCTGGAAGCCTTTCATTTTACAGCCCTGAGAGAGTCTCT 960
Oy      961 TTGCTAGTTGAATTAATGTGTGTGTTTTCCTGTAATAGCAAAATTAATTAATTAATG 1020
Db      961 TTGCTAGTTGAATTAATGTGTGTGTTTTCCTGTAATAGCAAAATTAATTAATTAATG 1020
Oy      1021 AAAAGTT 1027
Db      1021 AAAAGTT 1027

```

RESULT 4  
BC072396 1825 bp mRNA linear PRI 30-JUN-2004  
LOCUS Homo sapiens T cell receptor gamma variable 9, mRNA (cDNA clone MGC:90486 IMAGE:5755352), complete cds.

ACCESSION BC072396  
VERSION BC072396.1 GI:47939858  
KEYWORDS MGC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1825)  
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton T.E., Soares M.B., Bonaldo M.F., Caevari L., Scheetz T.E., Brownstein M.J., Uedin T.B., Toshiyuki S., Carinini P., Prange C., Kana S.S., Loquellano N.A., Peters G.J., Abremsen R.D., Mulihy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fehey J., Helton B., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shenchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Meyer R.M., Butlerfield Y.S., Krzywicki M.I., Skalka U., Smallus D.E., Scherch A., Schein J.E., Jones S.J. and Marra M.A.

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
PUBMED 12477932  
REFERENCE 2 (bases 1 to 1825)  
AUTHORS Strausberg R.  
TITLE Direct Submission  
JOURNAL Submitted (01-JUN-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer

## REMARK COMMENT

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>  
Contact: MGC help desk  
Email: [gcgaps-remail.nih.gov](mailto:gcgaps-remail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc\\_mgc@nsl.nih.gov](mailto:nisc_mgc@nsl.nih.gov)  
Ahter, N., Ayle, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakeley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Grant, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R., Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantirpop, S., Thomas, P.J., Touchman, J.W., Teurgeson, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
Series: IRAC Plate: 1/4 Row: c Column: 15  
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

## FEATURES SOURCE

Location/Qualifiers  
1..1825  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:90486 IMAGE:5755352"  
/issue\_type="Blood adult leukocytes"  
/clone\_id="NIH\_MGC\_118"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
1..1825  
/gene="TRGV9"  
/note="synonyms: TCRGV9, V2"  
/db\_xref="GeneID:6983"  
/db\_xref="IMGT/IRGM:TRGV9"  
378..1325  
/gene="TRGV9"  
/codon\_start=1  
/product="TRGV9 protein"  
/protein\_id="AAH72396.1"  
/db\_xref="GI:47939859"  
/db\_xref="GeneID:6983"  
/translaltion="MLSLHSTAVVAGLVCVAGHLEPOISTSTKLSTKLBCV VSGTIGATSVVYWRERGVSTQVIGVSGVQYRKSGIPSGFRVDRIPSTSLT TTHNVEKODIATYICALMEVFGELGKIKRFGPKLITDKDLADVSPPTFLPS IAEKSLQKAGTYICLLEFPDVIKIHMERKSTLIGSDEGNTMKNDYMKPSWLT VPEKSLDKERCIYRHNKNGVDQEIIFPIKTDVTMPKDCSDANDTLLOLT NTSAYVYVYLLLLKSVVYPAITCTCLARRAFCCNGSKS"

## CDS

## ORIGIN

Query Match 98.3%; Score 1010; DB 8; Length 1825;  
Best local similarity 99.5%; Pred. No. 8e-249;  
Matches 1013; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Oy 7 GAGTTGGGCAAAAATCAAGTATTTGGTCCCGGAACAAAGCTTACATGAGATTA 66  
Db 750 GAGTTGGGCAAAAATCAAGTATTTGGTCCCGGAACAAAGCTTATCATGAGATTA 809  
Oy 67 CAACCTGATGAGATGTTTCCCAAGCCCACTATTTTCTTCTCTCAATGCTGAACA 126  
Db 810 CAACCTGATGAGATGTTTCCCAAGCCCACTATTTTCTTCTCTCAATGCTGAACA 869  
Oy 127 AAGCTCAGAAAGCTGGAACATACCTTTGCTCTTGAAGAAATTTTCCCTGATGATT 186

```

Db      870 AAGTCAGAGGCGTGACATACCTTGTCTTCTGAGAAATTTTCCCTGATGTTATT 929
Qy      187 AAGTATCATTTGGCAAGAAAAGCAACCAATTTCTGGATTCCTCCAGAGGGGAAACC 246
Db      930 AAGTATCATTTGGCAAGAAAAGCAACCAATTTCTGGATTCCTCCAGAGGGGAAACC 989
Qy      247 ATGAAGACTTAACGACATCATGAAATTTAGTGTGTTAAACGGTGCAGAAAAGTCACTG 306
Db      990 ATGAGACTAATGACATCATGAAATTTAGTGTGTTAAACGGTGCAGAAAAGTCACTG 1049
Qy      307 GACAAAGACACAGATGATGTCAGACATGAGATATAATAAAACGAGTTGATCAAGAA 366
Db      1050 GACAAAGACACAGATGATGTCAGACATGAGATATAATAAAACGAGTTGATCAAGAA 1109
Qy      367 ATATATCTTCTCTCAATTAAGACGAGATGTCATCAATGATGCCAAAGACATTTGTTCA 426
Db      1110 ATATATCTTCTCTCAATTAAGACGAGATGTCATCAATGATGCCAAAGACATTTGTTCA 1169
Qy      427 AAGATGCAATGATGATGTCAGACATGAGATATAATAAAACGAGTTGATCAAGAA 486
Db      1170 AAGATGCAATGATGATGTCAGACATGAGATATAATAAAACGAGTTGATCAAGAA 1229
Qy      487 CTCCTCTGCTCTCTCAAGAGTGTGTATTTTTCATCATCACCTGCTGTCTGCTTGA 546
Db      1230 CTCCTCTGCTCTCTCAAGAGTGTGTATTTTTCATCATCACCTGCTGTCTGCTTGA 1289
Qy      547 AGAAGCGCTTTCTGCTGATGAGAGAAATCATTAAGACGCTGGCAAGAGGCGCAT 606
Db      1290 AGAAGCGCTTTCTGCTGATGAGAGAAATCATTAAGACGCTGGCAAGAGGCGCAT 1349
Qy      607 CTTTTCCTCATCGGTTATTTGCTCTAGAGAGGCTTCTGAGGATCTAGTGGGCTTCTT 666
Db      1350 CTTTTCCTCATCGGTTATTTGCTCTAGAGAGGCTTCTGAGGATCTAGTGGGCTTCTT 1409
Qy      667 TCTGGTGTGGGCGCATTTCTAGTGTCTAGTGTGACTATTTCTATCATTTATGTAACGG 726
Db      1410 TCTGGTGTGGGCGCATTTCTAGTGTGACTATTTCTATCATTTATGTAACGG 1469
Qy      727 TTTTCAAAACGAGTGGGCAACAGAGAACTCTCATCTCTGTAATACATGAGGATAGCCAC 786
Db      1470 TTTTCAAAACGAGTGGGCAACAGAGAACTCTCATCTCTGTAATACATGAGGATAGCCAC 1529
Qy      787 GGGGATCTCCAGACCAATCTCTCATGTTTTCACAGCTCTCCAGCAACCCCAATAG 846
Db      1530 GGGGATCTCCAGACCAATCTCTCATGTTTTCACAGCTCTCCAGCAACCCCAATAG 1589
Qy      847 CGCCTGTAATAGTGTAGACATCTGCGGCTTCTAGCCTTGTCCCTCTCTTATGTTCTT 906
Db      1590 CGCCTGTAATAGTGTAGACATCTGCGGCTTCTAGCCTTGTCCCTCTCTTATGTTCTT 1649
Qy      907 AATCAGATTAATGCTGCTGAGAGCTTTTCAATTTTACAGCCCTGAGACAGTCTTTTCTTA 966
Db      1650 AATCAGATTAATGCTGCTGAGAGCTTTTCAATTTTACAGCCCTGAGACAGTCTTTTCTTA 1709
Qy      967 GTTGAATTAATGCTGCTGTTTTCGTAATTAAGAAAATAATTAATAAAATGAAAA 1024
Db      1710 GTTGAATTAATGCTGCTGTTTTCGTAATTAAGAAAATAATTAATAAAATGAAAA 1767

```

```

RESULT 5
BC062761      1054 bp  mRNA  linear  PRI 16-DEC-2003
DEFINITION   Homo sapiens T cell receptor gamma variable 9, mRNA (cDNA clone
IMAGE:4248993), partial cds.
ACCESSION    BC062761
VERSION      BC062761.1  GI:38540986
KEYWORDS
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominae; Homo.
REFERENCE    1 (bases 1 to 1054)

```

```

AUTHORS
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, K.H., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Burow, K.H., Scheaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Mak, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Uedl, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huiyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E.,
Scherer, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1054)
REFERENCE
Strausberg, R.
Direct Submission
Submitted (25-NOV-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,
USA
NIH-MGC Project URL: http://mgc.ncl.nih.gov
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: CLONTECH
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
infobogsc.bc.ca
Steven Jones, Jennifer Amano, Ian Bosdet, Yaron Butterfield,
Susana Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeay, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabh, Parvaneh Saeedi, Jacqueline
Schein, Duane Smalins, Michael Smith, Lorraine Spence, Jeff Sprott,
Michael Thorne, Miranada Tsai, Natsja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: IRAL Plate: 51 Row: e Column: 16
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis.
location/Qualifiers
1. 1054
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4248993"
/cisue_type="Prostate"
/clone_lib="NIH MGC_83"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
<1. 1054
/feature="TRGV9"
/note="synonyms: TCRGV9, V2"
/db_xref="GeneID:6983"
<1. 577
/feature="TRGV9"
/codon_start=2
/product="TRGV9 protein"
/protein_id="AAH62761.1"

```

```
/db_xref="GI:38540987"
/db_xref="GeneID:6983"
/db_xref="IMG/LIGM:TRGV9"
/translation="GGGKIKVFGPGTULITDKOLADVSPKPTIFPSIAETKLOK
AGTLCLEKEFPVIRIKHMOEKSNITLSDGNTKNTWKPFMLVPEKSLDK
EHRCTVRHNNKNGVDSEIIFPPIKTIVITTDPPKNSKDAKDILLQLNTSAYMY
LILLKSVVFAIITCCLRTAFCCNGEKS"
74..364
/misc_feature
/gene="TRGV9"
/note="IGC; Region: Immunoglobulin domain constant region
subfamily"
/db_xref="CDD:c000098"

ORIGIN
Query Match      98.2%; Score 1008.8; DB 8; Length 1054;
Best Local Similarity 99.8%; Pred. No. 1.7e-248;
Matches 1010; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 16 AAAAAAAAAAGGATTTTGGTCCCGAAACAAAGCTTATCATTAACAGATTAACAACTTGAT 75
DB 11 AAAAAAAAAAGGATTTTGGTCCCGAAACAAAGCTTATCATTAACAGATTAACAACTTGAT 70
OY 76 GCAGATGTTCCCGAAGCCACTATTTTCTCCTCAATTGCTGAACAAAGCTCCAG 135
DB 71 GCAGATGTTCCCGAAGCCACTATTTTCTCCTCAATTGCTGAACAAAGCTCCAG 130
OY 136 AAGGCTGGAACATACCTTGTCTCTTGAGAAATTTTCCCTGATGTTATTAGATACAT 195
DB 131 AAGGCTGGAACATACCTTGTCTCTTGAGAAATTTTCCCTGATGTTATTAGATACAT 190
OY 196 TGGCAAGAAAAGAGACAACGATTTGGGATCCAGAGGGGAGACAACATGAGACT 255
DB 191 TGGCAAGAAAAGAGACAACGATTTGGGATCCAGAGGGGAGACAACATGAGACT 250
OY 256 AAGCAACATCATGAATTTAGCTGTTAAGCGGCGCAGAAAAGTCACTGGAACAAAGAA 315
DB 251 AAGCAACATCATGAATTTAGCTGTTAAGCGGCGCAGAAAAGTCACTGGAACAAAGAA 310
OY 316 CACAGATGATCTGCAGACATGAGATATATATAACGAGATTGATCAAGAAATTAATCTTT 375
DB 311 CACAGATGATCTGCAGACATGAGATATATATAACGAGATTGATCAAGAAATTAATCTTT 370
OY 376 CCTCCATTAAGACGGATGTCTCAATGATGCCAAAGACAAATGTTTCAAAAGATGCA 435
DB 371 CCTCCATTAAGACGGATGTCTCAATGATGCCAAAGACAAATGTTTCAAAAGATGCA 430
OY 436 AATGATACCTAGCTGCTGCAAGCTCAACACCTGTGATATTAATGATACCTCTCTG 495
DB 431 AATGATACCTAGCTGCTGCAAGCTCAACACCTGTGATATTAATGATACCTCTCTG 490
OY 496 CTCTCAAGAGTGTGTCTATTTTGCATCATCACTGTCTGTCTTGAAGAACGGCT 555
DB 491 CTCTCAAGAGTGTGTCTATTTTGCATCATCACTGTCTGTCTTGAAGAACGGCT 550
OY 556 TTCTGCTCAATGAGAGAAATCATTAACAGCGGTGGCAAGAGAGCCATCTTTTCTC 615
DB 551 TTCTGCTCAATGAGAGAAATCATTAACAGCGGTGGCAAGAGAGCCATCTTTTCTC 610
OY 616 ATCGGTTATTTGCTCAGAGAGGCTCTTGAGAGATCTAGTTGGGCTTTCTTTCTGGGTTT 675
DB 611 ATCGGTTATTTGCTCAGAGAGGCTCTTGAGAGATCTAGTTGGGCTTTCTTTCTGGGTTT 670
OY 676 GGGCCATTTCAATTCTCATGTGTACTATTCTATCATTTATTTATTAACGGTTTCAAC 735
DB 671 GGGCCATTTCAATTCTCATGTGTACTATTCTATCATTTATTTATTAACGGTTTCAAC 730
OY 736 CAGTGGGACACAGAGAACTCACTGTATTAACATGAGAGATAGCCACGGGATCTC 795
DB 731 CAGTGGGACACAGAGAACTCACTGTATTAACATGAGAGATAGCCACGGGATCTC 790
OY 796 CAGCAACCAATCTCTCAAGTTTTCAGACGCTCTCCAGCAACCCAAATAGGCGCTGCTA 855
DB 791 CAGCAACCAATCTCTCAAGTTTTCAGACGCTCTCCAGCAACCCAAATAGGCGCTGCTA 850
```

```
OY 856 TAGTGAGACATCTGCGGCTTCTAGCCTTGTCCCTCTCTTAGTGTCTTTAATCAAGATA 915
DB 851 TAGTGAGACATCTGCGGCTTCTAGCCTTGTCCCTCTCTTAGTGTCTTTAATCAAGATA 910
OY 916 ACTGCTCGAAGCCTTTCATTTTACAGCCCTGAAGAGCTTCTTGTCTAGTTGAATTA 975
DB 911 ACTGCTCGAAGCCTTTCATTTTACAGCCCTGAAGAGCTTCTTGTCTAGTTGAATTA 970
OY 976 TGTGTTGTGTTTTTCCGTAATTAAGCAAAATTAATTTAAAAAATGAAAAGTT 1027
DB 971 TGTGTTGTGTTTTTCCGTAATTAAGCAAAATTAATTTAAAAAATGAAAAGTT 1022

RESULT 6
CS032596 1530 bp DNA Linear PAT 10-MAR-2005
LOCUS Sequence 2102 from Patent WO2005016962.
DEFINITION CS032596
ACCESSION CS032596.1 GI:60732060
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE
1 Abbas, A., Clark, H., Ouyang, W., Williams, M. P., Wood, W. I. and Wu, T. D.
Compositions and methods for the treatment of immune related
diseases
Patent: WO 2005016962-A 2102 24-FEB-2005;
Genentech, Inc. (US)
FEATURES
source
location/Qualifiers
1..1530
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match      95.8%; Score 984; DB 6; Length 1530;
Best Local Similarity 98.0%; Pred. No. 4e-242;
Matches 996; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

OY 12 GGGCAAAAAATCAAGGTATTTGGTCCCGAACAAGCTTATCATTAACAGATAACAAC 71
DB 515 GGGCAAAATTAAGAACTCTTGGCAGTGAACAACCTGTGTCAAGATTAACAAC 574
OY 72 TGATGCAAGTGTTCCTCCCAAGCCCACTATTTTCTTCTGATGCTGAAGAAAGCT 131
DB 575 TGATGCAAGTGTTCCTCCCAAGCCCACTATTTTCTTCTGATGCTGAAGAAAGCT 131
OY 132 CCAGAGGCTGGAACATACCTTGTCTCTTGAAGAAATTTTCCCTGATGTTATTAAGAT 191
DB 635 CCAGAGGCTGGAACATACCTTGTCTCTTGAAGAAATTTTCCCTGATGTTATTAAGAT 191
OY 192 ACATTTGCAAGAAAAGAGACCAACAGATTTCTGGATCCAGAGAGGGAAACCATGAA 251
DB 695 ACATTTGCAAGAAAAGAGACCAACAGATTTCTGGATCCAGAGAGGGAAACCATGAA 251
OY 252 GACTTAAGACACATACATTAAGTAAATTTAGCTGTTAAGCGTGCAGAAAAGTCACTGGA 311
DB 755 GACTTAAGACACATACATTAAGTAAATTTAGCTGTTAAGCGTGCAGAAAAGTCACTGGA 311
OY 312 AGAACAAGATGTATCGTCAGACATGAGATTAATTAATAACGAGATTGATCAAGAAATTA 371
DB 815 AGAACAAGATGTATCGTCAGACATGAGATTAATTAATAACGAGATTGATCAAGAAATTA 371
OY 372 CTTCCTCCCAATTAAGAGAGATGTATCAATAGATCCCAAGACAAATTTTCAAAAGA 431
DB 875 CTTCCTCCCAATTAAGAGAGATGTATCAATAGATCCCAAGACAAATTTTCAAAAGA 431
OY 432 TGCATAATACATCTAGCTGAGCTCAACAACCTCTGCAATTAATGATGATCTCT 491
```

Db 935 TGCMAATGATACACTACTGCTGAGCTCACAAACCTCTGATATTACATGATACCTCT 994  
Qy 492 CCGCTCCCTCAAGAGTGTGTCTATTGTCATCATACCTGCTGTCTGTAGAAAC 551  
Db 995 CCGCTCCCTCAAGAGTGTGTCTATTGTCATCATACCTGCTGTCTGTAGAAAC 1054  
Qy 552 GGGCTTCTGCTGCAATGAGAGAAATCATTAACAGCGGTGACCAAGAGGACCTTTT 611  
Db 1055 GGGCTTCTGCTGCAATGAGAGAAATCATTAACAGCGGTGACCAAGAGGACCTTTT 1114  
Qy 612 CCGTATCGTATATGCTCTGTAAGAGCGTCTTGTAGAGATCTAGTGGGCTTTCTTTG 671  
Db 1115 CCGTATCGTATATGCTCTGTAAGAGCGTCTTGTAGAGATCTAGTGGGCTTTCTTTG 1174  
Qy 672 GTTTGGGCAATTCAGTCTCATGTGTACTATCTATCATTTATTTATTAACGGTTTC 731  
Db 1175 GTTTGGGCAATTCAGTCTCATGTGTACTATCTATCATTTATTTATTAACGGTTTC 1234  
Qy 732 AAACCAAGTGGGACACAGAGAACCTCACTGTATTAACAATGAGAAATGACGCGA 791  
Db 1235 AAACCAAGTGGGACACAGAGAACCTCACTGTATTAACAATGAGAAATGACGCGA 1294  
Qy 792 TCTCCAGACCAATCTCTCAATGTTTTCAAGCTCTCCAGCCCAACCAATAGCGCT 851  
Db 1295 TCTCCAGACCAATCTCTCAATGTTTTCAAGCTCTCCAGCCCAACCAATAGCGCT 1354  
Qy 852 GCTATAGTGAACATCTCGGCTTCTAGCTTCTCTCTTATGTTCTTTATCA 911  
Db 1355 GCTATAGTGAACATCTCGGCTTCTAGCTTCTCTCTTATGTTCTTTATCA 1414  
Qy 912 GATPACGCGTGAAGCTTCTTATTAACGCGCTGAGAGCTCTTTGCTAGTTGA 971  
Db 1415 GATPACGCGTGAAGCTTCTTATTAACGCGCTGAGAGCTCTTTGCTAGTTGA 1474  
Qy 972 ATTATGTGTGTGTTTTCCGTATTAAGCAAAATTAATTAATAAAGTT 1027  
Db 1475 ATTATGTGTGTGTTTTCCGTATTAAGCAAAATTAATTAATAAAGTT 1530

RESULT 7  
CS041548 1530 bp DNA linear PAT 22-MAR-2005  
LOCUS Sequence 2102 from Patent WO2005019258.  
DEFINITION CS041548  
ACCESSION CS041548.1 GI:61848985  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE  
AUTHORS 1  
TITLE Abbae, A., Clark, H., Ouyang, W., Williams, P. M., Wood, W. I. and Wu, T. D.  
JOURNAL Compositions and methods for the treatment of immune related  
GENE diseases  
Patent: WO 2005019258-A 2102 03-MAR-2005;  
Genentech, Inc. (US)  
FEATURES  
source location/Qualifiers  
1..1530  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 95.8%; Score 984; DB 6; Length 1530;  
Best Local Similarity 98.0%; Pred. No. 4e-242;  
Matches 996; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 12 GGGCAAAAATCAAGTATTGTTCCCGAACAAGCTTATCATTAACAGATAAACAAT 71  
Db 515 GGGCAAAAATCAAGTATTGTTCCCGAACAAGCTTATCATTAACAGATAAACAAT 574  
Qy 72 TGATGAGATGTTTCCCAAGCCCAATATTTTCTTCTTCAATGCTGAACAAGCT 131

Db 575 TGATGAGATGTTTCCCAAGCCCAATATTTTCTTCTTCAATGCTGAACAAGCT 634  
Qy 132 CCAAGAGCTGGAACATACCTTTGCTTCTTGAATAATTTTCCCTGATTTAAAGAT 191  
Db 635 CCAAGAGCTGGAACATACCTTTGCTTCTTGAATAATTTTCCCTGATTTAAAGAT 694  
Qy 192 ACATGGAAGAAAGAAAGAGCAACAGATTCCTGGGATCCAGAGGGGAAACCATGAA 251  
Db 695 ACATGGAAGAAAGAAAGAGCAACAGATTCCTGGGATCCAGAGGGGAAACCATGAA 754  
Qy 252 GACTAAGACACATACATGAATTTAGCTGTTTAAAGGTCAGAAAGTCACTGACAA 311  
Db 755 GACTAAGACACATACATGAATTTAGCTGTTTAAAGGTCAGAAAGTCACTGACAA 814  
Qy 312 AGAACACAGATGATGCTCAGACATGAATTAATAAAGGAGTTGATTAAGAAATAT 371  
Db 815 AGAACACAGATGATGCTCAGACATGAATTAATAAAGGAGTTGATTAAGAAATAT 874  
Qy 372 CTTTCTCCAAATTAAGCGATGTCATCAATGATCCCAAGACAAATTTGTTCAAGA 431  
Db 875 CTTTCTCCAAATTAAGCGATGTCATCAATGATCCCAAGACAAATTTGTTCAAGA 934  
Qy 432 TGCMAATGATACACTACTGCTGAGCTCAAAACACTCTGATATTACATGATACCTCT 491  
Db 935 TGCMAATGATACACTACTGCTGAGCTCAAAACACTCTGATATTACATGATACCTCT 994  
Qy 492 CCGTCTCTCAAGAGTGTGTCTATTGTCATCATCACTGCTGTCTGTTAAGAAC 551  
Db 995 CCGTCTCTCAAGAGTGTGTCTATTGTCATCATCACTGCTGTCTGTTAAGAAC 1054  
Qy 552 GGGCTTCTGCTGAATGAGAGAAATCATTAACAGAGGTGGCAAGAGGCGCATCTTT 611  
Db 1055 GGGCTTCTGCTGAATGAGAGAAATCATTAACAGAGGTGGCAAGAGGCGCATCTTT 1114  
Qy 612 CCGTATCGTATATGCTCTGTAAGAGCGTCTTGAAGATCTAGTGGGCTTTCTTCTG 671  
Db 1115 CCGTATCGTATATGCTCTGTAAGAGCGTCTTGAAGATCTAGTGGGCTTTCTTCTG 1174  
Qy 672 GTTTGGGCAATTCAGTCTCATGTGTACTATCTATCATTTATTTATTAACGGTTTC 731  
Db 1175 GTTTGGGCAATTCAGTCTCATGTGTACTATCTATCATTTATTTATTAACGGTTTC 1234  
Qy 732 AAACCAAGTGGGACACAGAGAACCTCACTGTATTAACAATGAGAAATGACGCGA 791  
Db 1235 AAACCAAGTGGGACACAGAGAACCTCACTGTATTAACAATGAGAAATGACGCGA 1294  
Qy 792 TCTCCAGACCAATCTCTCAATGTTTTCAAGCTCTCCAGCCCAACCAATAGCGCT 851  
Db 1295 TCTCCAGACCAATCTCTCAATGTTTTCAAGCTCTCCAGCCCAACCAATAGCGCT 1354  
Qy 852 GCTATAGTGAACATCTCGGCTTCTAGCTTCTCTCTTATGTTCTTTATCA 911  
Db 1355 GCTATAGTGAACATCTCGGCTTCTAGCTTCTCTCTTATGTTCTTTATCA 1414  
Qy 912 GATPACGCGTGAAGCTTCTTATTAACGCGCTGAGAGCTCTTTGCTAGTTGA 971  
Db 1415 GATPACGCGTGAAGCTTCTTATTAACGCGCTGAGAGCTCTTTGCTAGTTGA 1474  
Qy 972 ATTATGTGTGTGTTTTCCGTATTAAGCAAAATTAATTAATAAAGTT 1027  
Db 1475 ATTATGTGTGTGTTTTCCGTATTAAGCAAAATTAATTAATAAAGTT 1530

RESULT 8  
HUMTCRGA 1530 bp mRNA linear PRI 14-JAN-1995  
LOCUS Human T-cell receptor gamma chain VDJ-CII-CIII region mRNA,  
DEFINITION complete cds.  
ACCESSION M16768  
VERSION M16768.1 GI:339399  
KEYWORDS T-cell receptor.  
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.  
REFERENCE 1 (bases 1 to 1530)  
AUTHORS Krangel, M.S., Band, H., Hata, S., McLean, J. and Brenner, M.B.  
TITLE Structurally divergent human T cell receptor gamma proteins encoded  
by distinct C gamma genes  
JOURNAL Science 237 (4810), 64-67 (1987)  
PUBMED 2955517  
COMMENT Original source text: Human (cell line PBL Cl), cDNA to mRNA, clone  
PBIC1.15.  
Only a few V and J region genes occur within the genome. Diversity  
in T-cell receptors occur in the V-J recombination events. In some  
cases, this produces unproductively rearranged reading frames.  
Also, in some recombination events, some gene regions can be  
included more than one time creating even more diversity in TCR's.  
FEATURES  
Source  
Location/Qualifiers  
1..1530  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/map="7p15-p14"  
1..1530  
/gene="TCRG"  
<1..1530  
/gene="TCRG"  
/product="TCR-gamma mRNA"  
153..1085  
/gene="TCRG"  
/note="T-cell receptor (V-J-C) precursor"  
/codon\_start=1  
/protein\_id="AA61110.1"  
/db\_xref="GI:339400"  
/db\_xref="GDB:G00-120-407"  
/translation="MLSLSHASTLAVALCVAGAGLLEPOPISTSTSTSLTARLCV  
VSGTISATSVTRERREPVYQPLVSIIDGIVKSGSIPSGKFEVDRIPEISTSL  
TINAVEKQDIAIVYCALLEGYKKLFGSGTLLVVDQLDADVPKPTIFLPSIAERK  
LQKAGTYCLLEKFPDVIKIHQEKSNITLGSQSGNTMTNDTYMKFSLTVPERK  
LDKHCRCIVRHNNKNGVDQIIIPPIKTDIVTMDPKDNCSDANDTLQLTNTSAY  
YMYLLKLSKVVYFAIITCCLRRTPFCNENKS"  
153..197  
/gene="TCRG"  
/note="T-cell receptor signal peptide"  
198..1082  
/gene="TCRG"  
/product="T-cell receptor (V-J-CI-CII-CIII)"  
517..518  
/gene="TCRG"  
/organism="Homo sapiens"  
564..565  
/gene="TCRG"  
/organism="Homo sapiens"  
894..895  
/gene="TCRG"  
/organism="Homo sapiens"  
942..943  
/gene="TCRG"  
/organism="Homo sapiens"  
ORIGIN Chromosome 7p15.  
Query Match 95.8%; Score 984; DB 8; Length 1530;  
Best Local Similarity 98.0%; Pred. No. 4e-242;  
Matches 996; Conservative 0; Mismatches 20; Indels 0; Gaps 0;  
QY 12 GGGCAAAAATCAAGATTTGGTCCCGGAACAAGCTTATCATTAAGATAAACAAT 71  
DB 515 GGGAAATTATAAGAACTCTTGGCAAGTGAACAACACTGTGTGCACAGATAAACAAT 574  
QY 72 TGATGACAGATGTTTCCCAAGCCACCATATTTTCTTCCTCAATTGCTGAAACAAGCT 131  
DB 575 TGAATGACATGTTTCCCAAGCCACCATATTTTCTTCCTCAATTGCTGAAACAAGCT 634

QY 132 CCAAGAGCTGGAACATACCTTTGTCTTTGAGAAATTTTCCCTGATGTTATAGAT 191  
DB 635 CCAAGAGCTGGAACATACCTTTGTCTTTGAGAAATTTTCCCTGATGTTATAGAT 694  
QY 192 ACAATGGCAAGAAAGAGCAACAGATTTCTGGATCCCAAGAGGGAACCATGAA 251  
DB 695 ACATTGGCAAGAAAGAGCAACAGATTTCTGGATCCCAAGAGGGAACCATGAA 754  
QY 252 GACTAAGACATACATGAAATTTAGCTGTTAACGGTCCAGAAAGTCACTGACAA 311  
DB 755 GACTAAGACATACATGAAATTTAGCTGTTAACGGTCCAGAAAGTCACTGACAA 814  
QY 312 AGAACACAGATGATGTCGACATGAGATTAATTAAGAGTTGATCAAGAAATTA 371  
DB 815 AGAACACAGATGATGTCGACATGAGATTAATTAAGAGTTGATCAAGAAATTA 874  
QY 372 CTTTCCCTCAATTAAGACGATGTCATCAATGATCCCAAGACAAATGTTCAAGA 431  
DB 875 CTTTCCCTCAATTAAGACGATGTCATCAATGATCCCAAGACAAATGTTCAAGA 934  
QY 432 TGCAGATGATACCTACTGCTGACGCTCAACAAACCTCTGCATATTACATGTA 491  
DB 935 TGCAATGATACCTACTGCTGACGCTCAACAAACCTCTGCATATTACATGTA 494  
QY 492 CCGCTCTCAAGAGTGTGTTATTTTGGCATCATCACTGCTGTCTGTAGAGAAC 551  
DB 995 CCGCTCTCAAGAGTGTGTTATTTTGGCATCATCACTGCTGTCTGTAGAGAAC 1054  
QY 552 GCGTTTCTGTCGATGAGAGAAATCATTAACAGACGCTGAGCAAGAGGACCATCT 611  
DB 1055 GCGTTTCTGTCGATGAGAGAAATCATTAACAGACGCTGAGCAAGAGGACCATCT 1114  
QY 612 CCTCATGCTTATTTGCTCCTAGAAAGCTCTTGAAGATCTAGTTGGGCTTTCTTG 671  
DB 1115 CCTCATGCTTATTTGCTCCTAGAAAGCTCTTGAAGATCTAGTTGGGCTTTCTTG 1174  
QY 672 GTTTGGGCAATTCAGTTCTCAGTGTGTAATTCATTAATTTGATTAACGTTTC 731  
DB 1175 GTTTGGGCAATTCAGTTCTCAGTGTGTAATTCATTAATTTGATTAACGTTTC 1234  
QY 732 AAACCAATGGGACACAGAGAACTCACTCTGTAAATCAATGAGAAATGACAGG 791  
DB 1235 AAACCAATGGGACACAGAGAACTCACTCTGTAAATCAATGAGAAATGACAGG 1294  
QY 792 TCTCCAGACCAATCTCTCATGTTTTCACAGCTCCTCCAGCCAAACCAATAG 851  
DB 1295 TCTCCAGACCAATCTCTCATGTTTTCACAGCTCCTCCAGCCAAACCAATAG 1354  
QY 852 GCTATAGTGAACATCTCGGCTTCTAGCTTGTCCCTCTTATAGTGTCTTAATCA 911  
DB 1355 GCTATAGTGAACATCTCGGCTTCTAGCTTGTCCCTCTTATAGTGTCTTAATCA 1414  
QY 912 GATACTGCTGGAAGCTTTCATTTTACAGCCCTGAAGAGCTCTTGTGATGTA 971  
DB 1415 GATACTGCTGGAAGCTTTCATTTTACAGCCCTGAAGAGCTCTTGTGATGTA 1474  
QY 972 ATTATGATGATGTTTTCGTAATTAAGCAAAATTAATTAAGAAAGT 1027  
DB 1475 ATTATGATGATGTTTTCGTAATTAAGCAAAATTAATTAAGAAAGT 1530  
RESULT 9  
BC072387 1470 bp mRNA linear PRI 30-JUN-2004  
LOCUS Homo sapiens T cell receptor gamma variable 9, mRNA (CDNA clone  
DEFINITION IMAGE:5587705), partial cds.  
ACCESSION BC072387  
VERSION BC072387.1 GI:47938917  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

REFERENCE  
AUTHORS  
Hominidae; Homo.  
1 (bases 1 to 1470)  
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L.,  
Schetz T.E., Brownstein M.J., Udell T.B., Toshiyuki S.,  
Cernici P., Prange C., Raha S.S., Loquellano N.A., Peters G.J.,  
Abramson R.D., Mullany S.J., Boeck S.A., McEwan P.J.,  
McKernan K.J., Malek J.A., Guarnatone P.H., Richards S.,  
Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.M.,  
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S.,  
Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y.,  
Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D.,  
Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
Butterfield Y.S., Krzywinski M.I., Skalska U., Smailus D.E.,  
Schnerch A., Schein J.E., Jones S.J., Skalska U., Smailus D.E.,  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
12477932  
2 (bases 1 to 1470)  
Strausberg R.  
Direct Submission  
Submitted (01-JUN-2004) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [gcgaps-remail.nih.gov](mailto:gcgaps-remail.nih.gov)  
Tissue Procurement: Invitrogen  
cDNA Library Preparation: Life Technologies, Inc.  
Genome Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc.mgc@nih.gov](mailto:nisc.mgc@nih.gov)  
Blakesley R.W., Bouffard G.G., Breen K., Brinkley C., Brooks S.,  
Dietrich N.L., Granite S., Guan X., Gupta J., Haghighi P.,  
Hansen N., Ho S.-L., Karlins E., Kwong P., Latic P., Legaspi R.,  
Maduro Q.L., Mastello C., Maskeri B., Mastrian S.D., McCloskey J.C.,  
McDowell J., Pearson R., Stantiripod S., Thomas P.J., Touchman J.W.,  
Toung C., Vogt J.L., Walker M.A., Wetherby K.D., Wiggins L.,  
Young A., Zhang L.-H. and Green E.D.  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
Series: IRAC Plate: 174 Row: a Column: 5  
This clone was selected for full length sequencing because it  
passed the following selection criteria: Hexamer frequency ORF  
analysis.  
Location/Qualifiers  
1. 1470  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5587705"  
/tissue\_type="Ovary", pooled from 3 adults"  
/clone\_id="N1H\_MGC\_125"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
<1. 1470  
/gene="TRGV9"  
/note="synonyms: TCRGV9, VZ"  
/db\_xref="GeneID:6983"  
/db\_xref="IMGT/LIGM:TRGV9"  
<1. 899

ORIGIN  
Query Match 94.7%; Score 973; DB 8; Length 1470;  
Best Local Similarity 98.0%; Pred. No. 2.7e-239;  
Matches 985; Conservative 0; Mismatches 20; Indels 0; Gaps 0;  
/gene="TRGV9"  
/codon\_start=3  
/product="TRGV9 protein"  
/protein\_id="AAH72387.1"  
/db\_xref="GI:47938918"  
/db\_xref="GeneID:6983"  
/db\_xref="IMGT/LIGM:TRGV9"  
/translation="CYGAGHLEQPOISSTKLSKARLECVVSGITISATVYMYRE  
RGEVLPYVYKYLPGSGTLLVYDRLDADVSKRPTIPSPSTLTHVKKQDIATVYCA  
KFPDVYKHWQKSKNTLLGSGQNTMTKNDYKFSVLTPEKSLDKHRCIVRIE  
MNMNGVDQRIIPPIPTDVIYMDPKNCKSKNDNDLLDLTNSAYMYLLLLKSVV  
YFAITLCCLRRTAFPCNBEKS"  
20 AATCAAGGATTTGGTCCCGAACAAGCTTATCAATTAAGATAAACAATTGATGACAG 79  
337 ATAGAACTCTTTGGACGTGAAACAACCTTTGTCAAGATTAACAATTGATGACAG 396  
80 ATGTTCCCAAGCCCACTATTTTCTTCTTCAATTGCTGAACAAGCTCCAGAAG 139  
397 ATGTTCCCAAGCCCACTATTTTCTTCTTCAATTGCTGAACAAGCTCCAGAAG 456  
140 CTGGAACATACCTTTGCTTTCTTGAAGAAATTTTCCCTGATGTTATTAATACATTGGC 199  
457 CTGGAACATACCTTTGCTTTCTTGAAGAAATTTTCCCTGATGTTATTAATACATTGGC 516  
200 AAGAAAAGAACCAACAGATTTCTGGGATCCGAGAGGGGAAACCATGAAGCTAAG 259  
517 AAGAAAAGAACCAACAGATTTCTGGGATCCGAGAGGGGAAACCATGAAGCTAAG 576  
260 ACAATACATGAATTTAGCTGTGTTAACGCTGCAGAAAAGTCACTGACAAAGAACACA 319  
577 ACAATACATGAATTTAGCTGTGTTAACGCTGCAGAAAAGTCACTGACAAAGAACACA 636  
320 GATGATCGTCAACATGAGATAATTAATAAAGGAGTTGATCAAGAAATATCTTCTTC 379  
637 GATGATCGTCAACATGAGATAATTAATAAAGGAGTTGATCAAGAAATATCTTCTTC 696  
380 CAATTAAGCGATGTCATCAAGATGATCCCAAGCAATTTGTTCAAAAGATGAATG 439  
697 CAATTAAGCGATGTCATCAAGATGATCCCAAGCAATTTGTTCAAAAGATGAATG 756  
440 ATACACTACTGCTGACAGCTCAACAACCTCTGCATATTACATGATCCTCTGCTCC 499  
757 ATACACTACTGCTGACAGCTCAACAACCTCTGCATATTACATGATCCTCTGCTCC 816  
500 TCAAGATGTCATTTTGGCATCATCACTGCTGCTGCTTGAAGAACGGCTTTCT 559  
817 TCAAGATGTCATTTTGGCATCATCACTGCTGCTGCTTGAAGAACGGCTTTCT 876  
560 GCTGCATGAGAGAAATCAATAACAGCGTGGCAAGAGGCCATCTTCTCTATCG 619  
877 GCTGCATGAGAGAAATCAATAACAGCGTGGCAAGAGGCCATCTTCTCTATCG 936  
620 GTATTGTCCTTGAAGCGCTCTTCTGAGATCTAGTTGGCTTTCTTCTGGTTGGGC 679  
937 GTATTGTCCTTGAAGCGCTCTTCTGAGATCTAGTTGGCTTTCTTCTGGTTGGGC 996  
680 CATTTAGTTCTATGTGTACTATTCTATCTATTATGTAATACGGTTTCAACACAGT 739  
997 CATTTAGTTCTATGTGTACTATTCTATCTATTATGTAATACGGTTTCAACACAGT 1056  
740 GGGCAGACAGAAACCTCACTGTAATAACATGAGGAATAGCAAGGAGATCTCCAGC 799  
1057 GGGCAGACAGAAACCTCACTGTAATAACATGAGGAATAGCAAGGAGATCTCCAGC 1116  
800 ACCAATCTCCATGTTTTCACAGCTCTCCAGCAACCAATAGCGCTCTATAGT 859

Db 1117 ACACATCTCTCAGATGTTTCCACAGCTCTCTCAGCCAAACCAATAGGCGCTGCTATAGT 1176  
Qy 860 GTAGACATCTGCGGCTTCTAGCTTGTCCCTCTCTTATAGTCTTTATCAGATTAATCTG 919  
Db 1177 GTAGACATCTGCGGCTTCTAGCTTGTCCCTCTCTTATAGTCTTTATCAGATTAATCTG 1236  
Qy 920 CCTGGAAGCTTTTCACTTTTACACGCGCTGGAAGCAGTCTTCTTGTGATTAATTAATG 979  
Db 1237 CCTGGAAGCTTTTCACTTTTACACGCGCTGGAAGCAGTCTTCTTGTGATTAATTAATG 1296  
Qy 980 GTGTGTTTTTCCGTATATAGCAAAATTAATTTAAATAATGAAAA 1024  
Db 1297 GTGTGTTTTTCCGTATATAGCAAAATTAATTTAAATAATGAAA 1341  
RESULT 10  
HUMTCRGAAD 958 bp mRNA linear PRI 14-JUN-1995  
LOCUS Homo sapiens (clone HGP02) T cell receptor gamma-chain mRNA, Cl  
DEFINITION Homo sapiens (clone HGP02) T cell receptor gamma-chain mRNA, Cl  
ACCESSION M27334 GI:540463  
VERSION M27334.1  
KEYWORDS T cell receptor gamma-chain; constant region.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.  
1 (bases 1 to 958)  
Yoshikai, Y., Toyonaga, B., Koga, Y., Kimura, N., Griesse, H. and  
Mak, T. W.  
Repetitive of the human T cell gamma genes: high frequency of  
nonfunctional transcripts in thymus and mature T cells  
Eur. J. Immunol. 17 (1), 119-126 (1987)  
2949984  
COMMENT On Sep 21, 1994 this sequence version replaced gi:341672.  
Original source text: Homo sapiens cDNA to mRNA.  
FEATURES  
SOURCE  
Location/Qualifiers  
1..958  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/map="7p15-p14"  
/clone="HGP02"  
/cell\_type="PHA-stimulated peripheral T cell"  
1..958  
/gene="TCRCG1"  
<1..522  
/gene="TCRCG1"  
/codon\_start=1  
/product="T cell receptor gamma chain"  
/protein\_id="AAA6114.1"  
/db\_xref="GI:540464"  
/db\_xref="GDB:G00-120-408"  
/translation="DQDLADVSPKPTFLPSIAETKIQKAGTYICLLEKPPVYIK  
HMEKSNITLIGSGNTKNDYTKMSWLVPEKSLDKHRCIVRHNNKNGVDQ  
IIFPIKTDVITMDPKDNCXKADNTLLQLNTSAYVTYLLLSKVVPYAIITCL  
LRRTAFCCNGEKS"  
941..946  
/gene="TCRCG1"  
/note="G00-120-408"  
958  
/gene="TCRCG1"  
/note="G00-120-408"  
ORIGIN  
Query Match 91.3%; Score 938; DB 8; Length 958;  
Best Local Similarity 99.3%; Pred. No. 2.8e-230;  
Matches 953; Conservative 0; Mismatches 5; Indels 2; Gaps 1;  
Qy 61 GATAACAACCTGATGAGAGATGTTTCCCAAGCCCAACATTTTCTTCAATGCT 120  
Db 1 GATAACAACCTGATGAGATGTTTCCCAAGCCCAACATTTTCTTCAATGCT 60

Qy 121 GAAACAAAGCTCCGAAGGCTGGAACATACCTTTGTCTTGTGAAATTTTCCCTGAT 180  
Db 61 GAAACAAAGCTCCGAAGGCTGGAACATACCTTTGTCTTGTGAAATTTTCCCTGAT 120  
Qy 181 GTTATTAAGTATACCTGGCAAGAAAAGAGCAACAGATTTCCGAGTCCAGAGAGG 240  
Db 121 GTTATTAAGTATACCTGGCAAGAAAAGAGCAACAGATTTCCGAGTCCAGAGAGG 180  
Qy 241 AACACATGAAGACTAAGACACATACATGAAATTTAGTGTGTTAACGGTCCAGAAA 300  
Db 181 AACACATGAAGACTAAGACACATACATGAAATTTAGTGTGTTAACGGTCCAGAAA 240  
Qy 301 TCACCTGACAAAGAACACACATGATGTCAGACATGAAATTAATAAAGAGTTGAT 360  
Db 241 TCACCTGACAAAGAACACACATGATGTCAGACATGAAATTAATAAAGAGTTGAT 300  
Qy 361 CAAGAAATTAATCTTCCCTCCATTAAGACGAGATGTCACATGATGCCAAGACAT 420  
Db 301 CAAGAAATTAATCTTCCCTCCATTAAGACGAGATGTCACATGATGCCAAGACAT 360  
Qy 421 TGTTCAAAAGATGCAAAATGATATACACTGCTGACGCTCAAAACACTGCTGATATTAC 480  
Db 361 TGTTCAAAAGATGCAAAATGATATACACTGCTGACGCTCAAAACACTGCTGATATTAC 420  
Qy 481 ATGTACTCTCTGCTGCTCTCAAGAGTGTGCTTATTTTGGCATCATCACTGCTGTG 540  
Db 421 ACGTACTCTCTGCTGCTCTCAAGAGTGTGCTTATTTTGGCATCATCACTGCTGTG 480  
Qy 541 CTTGAAAGAACGGCTTCTGCTGCAATGAGAGAAATCATTAACAGAGTGGCAAGAGA 600  
Db 481 CTTGAAAGAACGGCTTCTGCTGCAATGAGAGAAATCATTAACAGAGTGGCAAGAGA 540  
Qy 601 GGCCATCTTTTCTCATCGTATATGTCCTTGAAGAGCTCTTGAAGATTAAGTGGC 660  
Db 541 GGCCATCTTTTCTCATCGTATATGTCCTTGAAGAGCTCTTGAAGATTAAGTGGC 600  
Qy 661 TTTCTTCTGGGTTTGGGCCATTTCAAGTCTCAATGTGTACTATTTATATTGTA 720  
Db 601 TTTCTTCTGGGTTTGGGCCATTTCAAGTCTCAATGTGTACTATTTATTTGTA 660  
Qy 721 TTAAGGTTTTTCAACGAGTGGGACACAGAGAACCTGCTGATTAACAATGAGAGAT 780  
Db 659 TTAAGGTTTTTCAACGAGTGGGACACAGAGAACCTGCTGATTAACAATGAGAGAT 720  
Qy 781 AGCCACGGCGATCTCCAGACCAATCTCTCATGTTTCCACAGCTCTCCAGCAACC 840  
Db 719 AGCCACGGCGATCTCCAGACCAATCTCTCATGTTTCCACAGCTCTCCAGCAACC 780  
Qy 841 AATAGCGCTGCTATAGTGAACATCTCGGCTTACCTTGTCTCTTATAGT 900  
Db 779 AATAGCGCTGCTATAGTGAACATCTCGGCTTACCTTGTCTCTTATAGT 838  
Qy 901 TTTCTTATACGATTAACGCGCTGGAAGCTTTCAATTTTACAGCGCTGAGAGAGCTCT 960  
Db 839 TTTCTTATACGATTAACGCGCTGGAAGCTTTCAATTTTACAGCGCTGAGAGAGCTCT 898  
Qy 961 TTGCTAGTTGAATATGATGATGTTTCCGTATATAGCAAAATTAATTTAAATAATG 1020  
Db 899 TTGCTAGTTGAATATGATGATGTTTCCGTATATAGCAAAATTAATTTAAATAATG 958  
RESULT 11  
BC039725 1655 bp mRNA linear PRI 06-JUN-2005  
LOCUS BC039725  
DEFINITION Homo sapiens similar to T-cell receptor gamma chain V region  
complete cds.  
ACCESSION BC039725  
VERSION BC039725.1 GI:24980796  
KEYWORDS MGC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 1655)

Strausner, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buettow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heile, F., Datchenko, L., Marisina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Urdin, T.B., Toshlyuk, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettleman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywicki, M.T., Skalska, U., Smailus, D.E., Buchanan, A., Schein, J.E., Jones, S.V. and Marra, M.A.

Mammalian Gene Collection Program Team  
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL  
PUBMED

12477932

2 (bases 1 to 1655)

CONSRM  
TITLE

NIH MGC Project

Direct Submission

Submitted (12-NOV-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>

Contact: MGC help desk

Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Galtherburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: [nisc\\_mgc@ncl.nih.gov](mailto:nisc_mgc@ncl.nih.gov)

Akter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,

Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,

Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,

Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaas, R.,

Maduro, O.L., Mastello, C., Maaker, B., Mastrian, S.D., McCloskey, J.C.,

McDowell, J., Pearson, R., Stantirpoo, S., Thomas, P.J., Touchman, J.W.,

Turgeon, C., Vogt, L., Walker, M.A., Wetherby, K.D., Wiggins, L.,

Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/ILNI, at: <http://image.llnl.gov>

Series: IRAX Plate: 82 Row: f Column: 20.

Location/Qualifiers

1. 1655

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="MGC:47828 IMAGE:5227869"

/tissue\_type="Pancreas, Spleen, adult pooled"

/clone\_id="NIH\_MGC\_120"

/lab\_host="DH10B"

/note="Vector: pCMV-SPORT6"

1. 1655

/gene="LOC442670"

/db\_xref="GeneID:442670"

191. 1162

/gene="LOC442670"

/codon\_start=1

/product="hypothetical protein LOC442670"

## ORIGIN

Query Match 88.0%; Score 903.8; DB 8; Length 1655;  
Best Local Similarity 92.9%; Pred. No. 1,7e-221;  
Matches 978; Conservative 0; Mismatches 27; Indels 48; Gaps 1;

/protein\_id="AAH39725.1"  
/db\_xref="GI:24980797"  
/db\_xref="GeneID:442670"  
/translation="MQMAALVLAIPSPASOKSNLEGRKSVIRPOTSSAETTCOLA  
EESNGTIIHMYLHQEGAPORLDYRYSYNSKVLSESVSGKTYTASTRNRLILRN  
LLENFGVYCAATWGPDPYKLFSGSTLLVVDKQDAVSPRPTFLPSIAETKQK  
ACTYICLLEKFPDPIIKIMOEKSNITLIGSGNMTKNDYMKRSMITPEBSIDK  
EHCIVHNNKNGKIDOEIIPPIKIVTVPKOSYSKADAVITMDPKDWSKDN  
DTLLQLTNTSAVYTYLLLLLSKVYFAIITCCLLRATFCNGERS"

20 AATCAAGTATTGGTCCCGAACAAGCTTATCTACAGTAAACAAGTATGACG 79  
552 ATAAAGAACTCTTGGCAAGTGAACAACCTGTTGTCACAGTAAACAAGTATGACG 611  
80 ATGTTCCCGCAAGCCCACTATTCTTCTTCAATGCTGAACAAGCTCCAGAAG 139  
612 ATGTTCCCGCAAGCCCACTATTCTTCTTCAATGCTGAACAAGCTCCAGAAG 671  
140 CTGGAACATACCTTTGCTCTTCTTGAAGAAATTTTCCCTGATGTTATTAAGATACATTGGC 199  
672 CTGGAACATACCTTTGCTCTTCTTGAAGAAATTTTCCCAAGTATTTAAGATACATTGGC 721  
200 AAGAAAAGAAAGCAACAGATCTGGGATCCAGAGGGGAACACCATGAACATACG 259  
732 AAGAAAAGAAAGCAACAGATCTGGGATCCAGAGGGGAACACCATGAACATACG 791  
260 ACACATACATGAATTAAGTGTGTTAAGGTGCAAGAAAGTCACTGGACCAAGAACACA 319  
792 ACACATACATGAATTAAGTGTGTTAAGGTGCAAGAAAGTCACTGGACCAAGAACACA 861  
320 GATGATCCGCAAGACATGAGATATTAAGAAACGAGTTGATCAAGAAATATCTTCTC 379  
852 GATGATCCGCAAGACATGAGATATTAAGAAACGAGTTGATCAAGAAATATCTTCTC 911  
380 CAATTAAGC-----GG 391  
912 CAATTAAGCAGATGTCACACAGTGAATCCCAAGACAGTTATTCAAAAGATGCAATG 971  
392 ATGTCATCAATGATATCCCAAGACATGTTGTTCAAGATGCAAGATGATACCTCTGC 451  
972 ATGTCATCAATGATATCCCAAGACATGTTGTTCAAGATGCAAGATGATACCTCTGC 1031  
452 TGCAGCTCAACAAACACCTGTCATTTATTAAGTACCTCTCTGCTCTCAAGATGTGG 511  
1032 TGCAGCTCAACAAACACCTGTCATTTATTAAGTACCTCTCTGCTCTCAAGATGTGG 1091  
512 TCTATTTTGGCATCATCACTGCTGTCGTGTTAAGAAAGCGCTTTCTGCTCAATGAG 571  
1092 TCTATTTTGGCATCATCACTGCTGTCGTGTTAAGAAAGCGCTTTCTGCTCAATGAG 1151  
572 AGAATCATPAACAGAGGTGGCAACAGAGGCCATCTTTCCATCCGTTATTCGCC 631  
1152 AGAATCATPAACAGAGGTGGCAACAGAGGCCATCTTTCCATCCGTTATTCGCC 1211  
632 AGAAGGCTCTTGAAGATTAAGTATGATTAAGCGTTTCTGAGGTTGGGCCATTCAGTTCT 691  
1212 AGAAGGCTCTTGAAGATTAAGTATGATTAAGCGTTTCTGAGGTTGGGCCATTCAGTTCT 1271  
692 CATGTGTGATATTTATTTATTTATTTATTAAGCGTTTCTGAGGTTGGGCCATTCAGTTCT 751  
1272 CATGTGTGATATTTATTTATTTATTTATTTATTAAGCGTTTCTGAGGTTGGGCCATTCAGTTCT 1331  
752 AACCTCACTGTGATTAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAG 811  
1332 AACCTCACTGTGATTAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAG 1391  
812 ATGTTTCCACAGCTCTCTGAGGCAACCAATATGAGGCTGCTATGATGATGATGATGATGAT 871

Db 1392 ATGTTTCCACAGCTCCTCCAGCCAAATAGCCCTGTCTATAGTGTAGACAGCTG 1451  
Qy 872 CGGCTTACGCTTGTCCCTCTCTTGTAGTGTCTTAATCAGATTAAGCTCGAAGCTT 931  
Db 1452 CGGCTTACGCTTGTCCCTCTCTTGTAGTGTCTTAATCAGATTAAGCTCGAAGCTT 1511  
Qy 932 TCATTTTACAGCCCTGAAAGAGCTCTTGTAGTGTGAATTAATGAGTGTGTTTCC 991  
Db 1512 TCATTTTACAGCCCTGAAAGAGCTCTTGTAGTGTGAATTAATGAGTGTGTTTCC 1571  
Qy 992 GTAATAGCAAAATTAATTTAAAAAATGAAAA 1024  
Db 1572 GTATTAAGCAAAATTAATTTAAAAAATGAAAA 1604

RESULT 12  
CO896252 1421 bp DNA linear PAT 05-NOV-2004  
LOCUS CO896252  
DEFINITION Sequence 76 from Patent WO2004076614.  
ACCESSION CO896252  
VERSION CO896252.1 GI:55468101  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominoidea; Homo.  
REFERENCE 1  
AUTHORS Hinzmann, B., Dahl, E., Rosenthal, A., Specht, T., Schmitz, A.,  
Beckmann, G., Brueemendorf, T., Kluemmann, H., Roepcke, S., Hermann, K.,  
Kinzhou, L., Pilarsky, C. and Staub, E.  
TITLE Human nucleic acid sequences obtained from prostatic carcinomas  
JOURNAL Patenz: WO 2004076614-A 76 10-SEP-2004;  
Hinzmann, Bernd (DB); Dahl, Edgar (DB); Rosenthal, Andre (DB);  
Specht, Thomas (DB); Schmitz, Armin (DB)  
FEATURES  
source  
1. 1421  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 87.8%; Score 902.2; DB 6; Length 1421;  
Best Local Similarity 92.8%; Pred. No. 4,5e-221;  
Matches 977; Conservative 0; Mismatches 28; Indels 48; Gaps 1;  
Qy 20 AATCAAGATATTGTCCTCCGAAACAAGCTTATCATTAAGATTAACAATTGATGAG 79  
Db 369 AATAAGAACTCTTGGCAGTGGAAACAACCTGTTGTACAGATTAACAATTGATGAG 428  
Qy 80 ATGTTTCCCAAGCCACTATTTTCTTCTTCAATTGCTGAACAAGCTCCAGAAG 139  
Db 429 ATGTTTCCCAAGCCACTATTTTCTTCTTCAATTGCTGAACAAGCTCCAGAAG 488  
Qy 140 CTGAACATACCTTGTCTTCTGAGAAATTTTCCCGATGTTATTAAGATTAACATTGGC 199  
Db 489 CTGAACATACCTTGTCTTCTTCTGAGAAATTTTCCCGATGTTATTAAGATTAACATTGGC 548  
Qy 200 AAGAAAAGAGAGCAACAGATTCTGGATCCAGAGAGGGAACAACATGAAGACTAAG 259  
Db 549 AAGAAAAGAGAGCAACAGATTCTGGATCCAGAGAGGGAACAACATGAAGACTAAG 608  
Qy 260 ACACATACATGAATTAATGCTGTGAACGGTGCAGAAAGTCACTGGAACAAGACA 319  
Db 609 ACACATACATGAATTAATGCTGTGAACGGTGCAGAAAGTCACTGGAACAAGACA 668  
Qy 320 GATGATGTGTGAGATGAATTAATAAAGAGGTTGATCAAGAAATTAATCTTCTC 379  
Db 669 GATGATGTGTGAGATGAATTAATAAAGAGGTTGATCAAGAAATTAATCTTCTC 728  
Qy 380 CAATTAAGAC-----GG 391  
Db 729 CAATTAAGACAGATGTACACAGATGATCCCAAGACAGTTATTCAAAAGATGCAATG 788

Qy 392 ATGTGATCAAGATGATCCCAAGACAAATTGTTCAAAAAGATGCAAAATGATACACTAGC 451  
Db 789 ATGTGATCAAGATGATCCCAAGACAAATTGTTCAAAAAGATGCAAAATGATACACTAGC 848  
Qy 452 TGCAGCTCAAAACACCTGTGCATATTATTAACATGACCTCTCTGCTCTCTCAAGTGTG 511  
Db 849 TGCAGCTCAAAACACCTGTGCATATTATTAACATGACCTCTCTGCTCTCTCAAGTGTG 908  
Qy 512 TCTATTTGCAATCAATCACTGCTGTCTGCTTAAGAAAGACGGCTTTCTGCTCAATGAG 571  
Db 909 TCTATTTGCAATCAATCACTGCTGTCTGCTTAAGAAAGACGGCTTTCTGCTCAATGAG 968  
Qy 572 AGAATCATTAACAGCGTGGCAAGAGAGCCATCTTCTCATGCGTTATGTCCT 631  
Db 969 AGAATCATTAACAGCGTGGCAAGAGAGCCATCTTCTCATGCGTTATGTCCT 1028  
Qy 632 AGAAGCTCTTGAAGATTAAGTGGGCTTTCTTCTGAGTTTGGCCATTTCACTTCT 691  
Db 1029 AGAAGCTCTTGAAGATTAAGTGGGCTTTCTTCTGAGTTTGGCCATTTCACTTCT 1088  
Qy 692 CATGTGTGACTATTTCTATCATTAATGATTAACGGTTTCAAAACAGTGGCACACAGAG 751  
Db 1089 CATGTGTGACTATTTCTATCATTAATGATTAATGATTAACGGTTTCAAAACAGTGGCACACAGAG 1148  
Qy 752 AACTCAGCTCTGTAATTAACAATAGAGAACGACGGCGATCTCCAGACCAATCTCTCC 811  
Db 1149 AACTCAGCTCTGTAATTAACAATAGAGAACGACGGCGATCTCCAGACCAATCTCTCC 1208  
Qy 812 ATGTTTCCACAGCTCCTCCAGCCAAACCAATAGCCCTGTATAGTATGACATCTG 871  
Db 1209 ATGTTTCCACAGCTCCTCCAGCCAAACCAATAGCCCTGTATAGTATGACATCTG 1268  
Qy 872 CGGCTTACGCTTGTCCCTCTCTTGTAGTGTCTTAATCAAGATTAAGCTCGAAGCTT 931  
Db 1269 CGGCTTACGCTTGTCCCTCTCTTGTAGTGTCTTAATCAAGATTAAGCTCGAAGCTT 1328  
Qy 932 TCATTTTACAGCCCTGGAAGCAGCTCTTGTGAGTGAATTAATGAGTGTGTTTCC 991  
Db 1329 TCATTTTACAGCCCTGGAAGCAGCTCTTGTGAGTGAATTAATGAGTGTGTTTCC 1388  
Qy 992 GTAATAGCAAAATTAATTTAAAAAATGAAAA 1024  
Db 1389 GTATTAAGCAAAATTAATTTAAAAAATGAAAA 1421

RESULT 13  
AX332835 1421 bp DNA linear PAT 09-JAN-2002  
LOCUS AX332835  
DEFINITION Sequence 3344 from Patent WO0194629.  
ACCESSION AX332835  
VERSION AX332835.1 GI:18123469  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominoidea; Homo.  
REFERENCE 1  
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Andrews, G.,  
Horrisgan, S., Soppet, D.R. and Weaver, Z.  
TITLE Cancer gene determination and therapeutic screening using signature  
JOURNAL gene sets  
Patent: WO 0194629-A 3344 13-DEC-2001;  
Avalon Pharmaceuticals (US)  
FEATURES  
source  
1. 1421  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 87.8%; Score 902.2; DB 6; Length 1421;

Best Local Similarity 92.8%; Pred. No. 4.5e-221;  
Matches 977; Conservative 0; Mismatches 28; Indels 48; Gaps 1;

```

QY 20 AATCAAGGATTTGGTCCCGAACAAGCTTATCATTAAGATTAACAACCTTGATGACG 79
DB 369 ATAGAAGAACTCTTGGGAGTGGAACAACCTTGTTGTCACAGATTAACAACCTTGATGACG 428
QY 80 ATGTTTCCCGCAAGCCCACTATTTTCTGCTTCAATGCTGAAACCAAGCTCCAGAGG 139
DB 429 ATGTTTCCCGCAAGCCCACTATTTTCTGCTTCAATGCTGAAACCAAGCTCCAGAGG 488
QY 140 CTGGAACATACCTTTGCTTCTTGAGAAATTTTCCCTGATGTTATTAAGATCAATGGC 199
DB 489 CTGGAACATACCTTTGCTTCTTGAGAAATTTTCCAGATTAATTAAGATCAATGGC 548
QY 200 AAGAAAAGAGAGACACAGATTCCTGGGATCCAGAGGGGAAACACCATGAAGCTAACG 259
DB 549 AAGAAAAGAGAGACACAGATTCCTGGGATCCAGAGGGGAAACACCATGAAGCTAACG 608
QY 260 ACAATATCAAGAAATTTAGCTGTGTAACGGTGCAGAAAGTCACTGGAACAAGAACACA 319
DB 609 ACAATATCAAGAAATTTAGCTGTGTAACGGTGCAGAAAGTCACTGGAACAAGAACACA 668
QY 320 GATGTATCGTCAGCATGAGAAATTAATAAAGAGAGTTGATCAAGAAATTAATCTTCTC 379
DB 669 GATGTATCGTCAGCATGAGAAATTAATAAAGAGAGTTGATCAAGAAATTAATCTTCTC 728
QY 380 CAATTAAGAC-----GG 391
DB 729 CAATTAAGACAGATGTCCACAGTGTGATCCAAAGACAGTATTAATAAAGATGCAATG 788
QY 392 ATGTCAATCAAGATGATCCCAAGAACAAATGTTCAAAAGATGCAAAAGATGCACTATGC 451
DB 789 ATGTCAATCAAGATGATCCCAAGAACAAATGTTCAAAAGATGCAAAAGATGCACTATGC 848
QY 452 TGCAGCTCACAACAACCTCTGCAATTAATTAACATGTAATCTCTCTCTCAAGAGTGG 511
DB 849 TGCAGCTCACAACAACCTCTGCAATTAATTAACATGTAATCTCTCTCTCAAGAGTGG 908
QY 512 TCTATTTGGCATTGATCAGCTGCTGTCTGTAGAGAGAGCGCTTTGCTGCAATGAG 571
DB 909 TCTATTTGGCATTGATCAGCTGCTGTCTGTAGAGAGAGCGCTTTGCTGCAATGAG 968
QY 572 AGAAATATTAACAGAGCTGGGACACAGAGAGGCACTTTTCTCTATGCGTATTTGCTCT 631
DB 969 AGAAATATTAACAGAGCTGGGACACAGAGAGGCACTTTTCTCTATGCGTATTTGCTCT 1028
QY 632 AGAAGCGTCTTGAGGATCTAGTTGGGCTTTCTTCTGGGTTTGGGCAATTCAGTTCT 691
DB 1029 AGAAGCGTCTTGAGGATCTAGTTGGGCTTTCTTCTGGGTTTGGGCAATTCAGTTCT 1088
QY 692 CATGTGTGATCTATTTCTATTAATTTGATTAACGATTTTCAACAGTGAGGACACAGAG 751
DB 1089 CATGTGTGATCTATTTCTATTAATTTGATTAATGTTTCAACAGTGAGGACACAGAG 1148
QY 752 AACCTCACTCTGATTAACAATGAGAAATGCGGCACTCTCAGACCAATCTCTCC 811
DB 1149 AACCTCACTCTGATTAACAATGAGAAATGCGGCACTCTCAGACCAATCTCTCC 1208
QY 812 ATGTTTCCACAGCTCTCCAGCAACCAATGAGGCTGTAATGATGAGATCATGCTG 871
DB 1209 ATGTTTCCACAGCTCTCCAGCAACCAATGAGGCTGTAATGATGAGATCATGCTG 1268
QY 872 CGGCTTGAAGCTTGTCTCTCTTGAATGTTCTTAATCAAGATTAAGCTGGAAGCTT 931
DB 1269 CGGCTTGAAGCTTGTCTCTCTTGAATGTTCTTAATCAAGATTAAGCTGGAAGCTT 1328
QY 932 TCATTTTACAGCCCTGAGAGAGCTTTTGTCTAGTTGAATTAATGAGTGTGTTTCC 991
DB 1329 TCATTTTACAGCCCTGAGAGAGCTTTTGTCTAGTTGAATTAATGAGTGTGTTTCC 1388
QY 992 GTATTAAGCAAAATTAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1024

```

```

DB 1389 GTATTAAGCAAAATTAATTAATAAATAAATAAATAAATAAATAAATAAATAAATA 1421

RESULT 14
HUMTCGMX 1421 bp mRNA linear PRI 03-AUG-1993
LOCUS Human T-cell receptor aberrantly rearranged gamma-chain mRNA from cell line HPB-MLT.
DEFINITION M13231.1 GI:339168
ACCESSION M13231
VERSION C-region; J-region; T-cell receptor gamma chain; V-region; antigen receptor; processed gene; pseudogene.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1421)
Dialynas,D.P., Murte,C., Quentemous,T., Bosa,J.M., Leiden,J.M., Seidman,J.G. and Strominger,J.L. Cloning and sequence analysis of complementary DNA encoding an aberrantly rearranged human T-cell gamma chain Proc. Natl. Acad. Sci. U.S.A. 83 (8), 2619-2623 (1986)
JOURNAL PUBMED 3458221
COMMENT Original source text: Human T-cell line HPB-MLT, cDNA to mRNA, clones pT-gamma-[1,2].
Computer-readable sequence of [1] kindly provided by D.P.Dialynas, 31-OCT-1986.
An aberrant rearrangement at the V-J junction causes a frame shift that leads to peptide termination at the 'taa' codon at positions 370-372.
FEATURES
source
location/Qualifiers
1..1421
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
4..979
/feature="pseudo-TCG pept"
/pseudo
/codon_start=1
4..63
/feature="pseudo-TCG signal pept"
/pseudo
64..976
/feature="pseudo-TCG mature pept"
/pseudo
mat_peptide
ORIGIN 16 bp upstream of Apal site; chromosome 7.
Query Match 87.8%; Score 902.2; DB 8; Length 1421;
Best Local Similarity 92.8%; Pred. No. 4.5e-221;
Matches 977; Conservative 0; Mismatches 28; Indels 48; Gaps 1;

```

Qy 320 GATGATCGTCAGCATGAGATAATATAAAACGAGATTGATCAAGAAATTATCTTCTC 379  
 Db 669 GATGATCGTCAGCATGAGATAATATAAAACGAGATTGATCAAGAAATTATCTTCTC 728  
 Qy 380 CAATTAAGAC-----CG 391  
 Db 729 CAATTAAGACGAGATGTCACAGTGATCCCAAGAGATTATCAAAAGATGCAAAATG 788  
 Qy 392 ATGTCATCAAGATGATCCCAAGAGATTGATCAAAAGATGCAAAATGATACACTATGC 451  
 Db 789 ATGTCATCAAGATGATCCCAAGAGATTGATCAAAAGATGCAAAATGATACACTATGC 848  
 Qy 452 TGCAAGTCACAAACCTCTGCATATTATACATGATCTCTCTCTCTCTCAAGAGTGG 511  
 Db 849 TGCAAGTCACAAACCTCTGCATATTATACATGATCTCTCTCTCTCTCAAGAGTGG 908  
 Qy 512 TCTATTTTGGCATATCATCTGCTGCTCTGCTTAAGAGACGGCTTTCTGCTGCAATGGAG 571  
 Db 909 TCTATTTTGGCATATCATCTGCTGCTGCTGCTTAAGAGACGGCTTTCTGCTGCAATGGAG 968  
 Qy 572 AGAATTCATAACAGACGGTGGACACAAAGAGGCAATCTTTCTCATCGGATTATGCTCT 631  
 Db 969 AGAATTCATAACAGACGGTGGACACAAAGAGGCAATCTTTCTCATCGGATTATGCTCT 1028  
 Qy 632 AGAAGCGCTCTCTGAGATCTAGTTGGGCTTTCTCTGAGGTTGGGCAATTCAGTTCT 691  
 Db 1029 AGAAGCGCTCTCTGAGATCTAGTTGGGCTTTCTCTGAGGTTGGGCAATTCAGTTCT 1088  
 Qy 692 CATGTGTACTATTTCTATCTATTTATTAAGGTTTCAACACAGTGGGACACAGAG 751  
 Db 1089 CATGTGTACTATTTCTATCTATTTATTAAGGTTTCAACACAGTGGGACACAGAG 1148  
 Qy 752 AACCTACTCTGTAATAACAATGAGAGATGACACGGGATCTCCAGACCAATCTCTCC 811  
 Db 1149 AACCTACTCTGTAATAACAATGAGAGATGACACGGGATCTCCAGACCAATCTCTCC 1208  
 Qy 812 ATGTTTTCACAGCTCTCTCCAGCCAAATAGCGCTCTATAGTGAACATCTCTG 871  
 Db 1209 ATGTTTTCACAGCTCTCTCCAGCCAAATAGCGCTCTATAGTGAACATCTCTG 1268  
 Qy 872 CGGCTTCAAGCTTGTCCCTCTCTTGAAGTTCTTAATCAAGATACTGCTGGAAGCTT 931  
 Db 1269 CGGCTTCAAGCTTGTCCCTCTCTTGAAGTTCTTAATCAAGATACTGCTGGAAGCTT 1328  
 Qy 932 TCATTTCACAGCCCTGAAGAGTCTTCTTGGCTAGTGAATTAATGATGATGTTTCC 991  
 Db 1329 TCATTTCACAGCCCTGAAGAGTCTTCTTGGCTAGTGAATTAATGATGATGTTTCC 1388  
 Qy 992 GTAATTAAGCAAAATTAATTTAAAAAATGAAAA 1024  
 Db 1389 GTAATTAAGCAAAATTAATTTAAAAAATGAAAA 1421

RESULT 15  
 HUMTREGAAA 1013 bp mRNA linear PRI 14-JAN-1995  
 LOCUS Homo sapiens (clone HGP08) T cell receptor gamma-chain mRNA, C2  
 DEFINITION  
 region.  
 ACCESSION M27331.1 GI:540458  
 VERSION M27331.1  
 KEYWORDS T cell receptor gamma-chain; constant region.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.  
 Yashikuni, Y., Toyonaga, B., Koga, Y., Kimura, N., Griesser, H. and  
 Mak, T. W.  
 Repertoire of the human T cell gamma genes: high frequency of  
 nonfunctional transcripts in thymus and mature T cells  
 JOURNAL Eur. J. Immunol. 17 (1), 119-126 (1987)  
 PUBMED 2949984

COMMENT On Sep 21, 1994 this sequence version replaced gi:341669.  
 Original source text: Homo sapiens PHA-stimulated peripheral T cell  
 cDNA to mRNA.

FEATURES  
 source  
 1..1013  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /map="7p15-p14"  
 /clone="HGP08"  
 /tissue\_type="PHA-stimulated peripheral T cell"  
 1..1013  
 /gene="TCRCG2"  
 /note="This CDS feature is included to show the  
 translation of the corresponding C\_region. Presently  
 translation qualifiers on C\_region features are illegal."  
 /codon\_start=1  
 /protein\_id="AA61112.1"  
 /db\_xref="GI:540459"  
 /translation="DKQLADVSPKPTIFPISAEKTLQKAGTYLCLLEKFPDIKI  
 HMOBKSNITLIGSGEYTMKNDPMYKFSWLTVPESLDKHRCIVHNNKNDDE  
 IIFPIKTDVITDPTKYNKSDANDVITNDPKDMSKDANDTLILQLTNSATMYTL  
 LILKSVYFAIITICLLGRTPCCNGSKS"  
 <1..570  
 /gene="TCRCG2"  
 /product="T cell receptor gamma chain"  
 /note="G00-120-409"  
 989..994  
 /gene="TCRCG2"  
 /note="G00-120-409"  
 1013  
 /gene="TCRCG2"  
 /note="G00-120-409"

polya\_signal  
 polyA\_signal  
 1013  
 /gene="TCRCG2"  
 /note="G00-120-409"

ORIGIN  
 Query Match 85.4%; Score 877.4; DB 8; Length 1013;  
 Best Local Similarity 94.0%; Pred. No. 1,18-214;  
 Matches 954; Conservative 0; Mismatches 11; Indels 50; Gaps 2;  
 Qy 61 GATAAACAATTGATGCAAGATGTTTCCCAAGCCCACTATTTTCTTCTCAATTGCT 120  
 Db 1 GATAAACAATTGATGCAAGATGTTTCCCAAGCCCACTATTTTCTTCTCAATTGCT 60  
 Qy 121 GAAACAAGCTCCGAAAGGCTGGAACATACCTTTGTCTCTGGAATAATTTCCCTGAT 180  
 Db 61 GAAACAAGCTCCGAAAGGCTGGAACATACCTTTGTCTCTGGAATAATTTCCCTGAT 120  
 Qy 181 GTTATTAAGATACATTGGCAAGAAAGAGCAACAGATTCTGGATGCCAGAGGG 240  
 Db 121 ATTATTAAGATACATTGGCAAGAAAGAGCAACAGATTCTGGATGCCAGAGGG 180  
 Qy 241 AACACATGAAGACTTAACGACATACATGAATTTAGCTGTTAACGCTGCCAGAAAG 300  
 Db 181 AACACATGAAGACTTAACGACATACATGAATTTAGCTGTTAACGCTGCCAGAAAG 240  
 Qy 301 TCACATGACAAAGAACACAGATGATGTCGACATGAGATTAATAAAGAGGATTGAT 360  
 Db 241 TCACATGACAAAGAACACAGATGATGTCGACATGAGATTAATAAAGAGGATTGAT 300  
 Qy 361 CAAGAAATTATCTTCTCCATTAAGAC----- 389  
 Db 301 CAAGAAATTATCTTCTCCATTAAGAC----- 389  
 Qy 390 -----GATGTCATCAATGATGCCAAGAGCAATGTTCAAAAGAT 432  
 Db 361 TATTCAAAGATGCAAAATGATGTCATCAATGATGCCAAGAGCAATGTTCAAAAGAT 420  
 Qy 433 GCAATGATTAACATGCTGCTGACGCTCAACAACTCTGATATTACATGACTCTCTC 492  
 Db 421 GCAATGATTAACATGCTGCTGACGCTCAACAACTCTGATATTACATGACTCTCTC 480

```
QY 493 CTGCTCTCAAGAGTGTGTCTAATTTTGCATCATCACTGCTGTCTGTTAGAGAAGC 552
   |||||
Db 481 CTGCTCTCAAGAGTGTGTCTAATTTTGCATCATCACTGCTGTCTGTTAGAGAAGC 540
   |||||
QY 553 GCTTTCTGCTGCAATGAGAGAAATCAATACAGAGGTGGCACAGAGGCCATCTTTTC 612
   |||||
Db 541 GCTTTCTGCTGCAATGAGAGAAATCAATACAGAGGTGGCACAGAGGCCATCTTTTC 600
   |||||
QY 613 CTGATGGGTATATGTCCCTTAGAAGCGTCTCTGAGATCTAGTGGGCTTCTTCTGGG 672
   |||||
Db 601 CTGATGGGTATATGTCCCTTAGAAGCGTCTCTGAGATCTAGTGGGCTTCTTCTGGG 660
   |||||
QY 673 TTGGGGCATTTCAAGTCTCATGTGTACTATCTATCATATTATGTATAAGGTTTCA 732
   |||||
Db 661 TTGGGGCATTTCAAGTCTCATGTGTACTATCTATCATATTATGTATAAGGTTTCA 718
   |||||
QY 733 AACGATGGGCACACAGAGAACTCACTCTGTATAACAAATGAGAAATAGCCAGCGAT 792
   |||||
Db 719 AACGATGGGCACACAGAGAACTCACTCTGTATAACAAATGAGAAATAGCCAGCGAT 778
   |||||
QY 793 CTCGAGACCAATCTCTCCATGTTTTCACAGCTCTCCAGCCAAACCAATAGCGCTG 852
   |||||
Db 779 CTCGAGACCAATCTCTCCATGTTTTCACAGCTCTCCAGCCAAACCAATAGCGCTG 838
   |||||
QY 853 CTATAGTGTAGACATCTCTGCGGCTTCTAGCCTTGCCCTCTCTAGTGTCTTAAATCAG 912
   |||||
Db 839 CTATAGTGTAGACATCTCTGCGGCTTCTAGCCTTGCCCTCTCTAGTGTCTTAAATCAG 898
   |||||
QY 913 ATAACTGCTGGAAGCCTTTCATTTTACAGCCCTGAAGCAGTCTTCTTGTAGTGA 972
   |||||
Db 899 ATAACTGCTGGAAGCCTTTCATTTTACAGCCCTGAAGCAGTCTTCTTGTAGTGA 958
   |||||
QY 973 TTATGTGTGTGTTTTCCTTAATAGCAAAATTAATTAATAAATGAAGTT 1027
   |||||
Db 959 TTATGTGTGTGTTTTCCTTAATAGCAAAATTAATTAATAAATGAAGTT 1013
   |||||
```

Search completed: December 10, 2005, 22:14:55  
Job time : 5319 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus model

Run on: December 8, 2004, 06:11:46 ; Search time 1877 Seconds

(without alignments)  
1461.269 Million cell updates/sec

Title: US-10-031-158-14

Perfect score: 307  
Sequence: 1 MOWPSPPLFFPQLKQSS.....RYGKRRRATRFMDRRGTP 58

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=x1h  
-O=/cgn2\_1/USPRO.epool/US10031158/runat\_06122004\_082659\_15887/apd\_query.fasta\_1.199  
-DB=genEmb1 -QMT=faastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=pcpt -THR MAX=100 -THR MIN=0 -ALIGN=40 -MODE=LOCAL  
-OUTMT=ptio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10031158 @CGN\_1\_1\_2527 @runat\_06122004\_082659\_15887 -NCPU=6 -ICPU=3  
-NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database :

GenEmb1:\*  
1: gb\_ha:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_scs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	307	100.0	410	9	HUMTCGCD1
2	307	100.0	426	6	AX884816 Sequence
3	307	100.0	426	6	BD024426 Sequence
4	307	100.0	470	6	CQ131517 Sequence

5	307	100.0	470	6	CQ290455 Sequence
6	307	100.0	477	6	CQ132293 Sequence
7	307	100.0	477	6	CQ200006 Sequence
8	307	100.0	477	6	CQ291096 Sequence
9	307	100.0	539	6	CQ409917 Sequence
10	307	100.0	539	6	CQ496769 Sequence
11	307	100.0	723	9	AJ583012 Homo sapi
12	307	100.0	726	9	AJ583014 Homo sapi
13	307	100.0	825	6	189883 Sequence 18
14	307	100.0	825	6	189883 Sequence 18
15	307	100.0	958	9	HUMTCRGAAD
16	307	100.0	1027	6	AX074415
17	307	100.0	1027	6	AF151103 Homo sapi
18	307	100.0	1054	9	BC062761 Homo sapi
19	307	100.0	1316	6	CQ493215 Sequence
20	307	100.0	1316	6	CQ494607 Sequence
21	307	100.0	1470	9	BC072387 Homo sapi
22	307	100.0	1530	9	HUMTCRGA
23	307	100.0	1825	9	BC072396 Homo sapi
24	307	100.0	66558	2	AC130306 Homo sapi
25	307	100.0	140691	9	AF159056 Homo sapi
26	307	100.0	171816	9	AC006033 Homo sapi
27	304	99.0	400	9	HSTCELG1
28	304	99.0	486	6	CQ493135 Sequence
29	302	98.4	316	6	CQ144550 Sequence
30	302	98.4	316	6	CQ204403 Sequence
31	302	98.4	316	6	CQ302987 Sequence
32	298	97.1	569	6	CQ511627 Sequence
33	294	95.8	330	9	HUMTCGCG
34	294	95.8	635	9	HSTRGCG4
35	294	95.8	720	9	HSTRGCG10
36	294	95.8	1013	9	HUMTCRGA
37	294	95.8	1160	9	HSTRGCG4
38	294	95.8	1402	9	HUMTCRGA
39	294	95.8	1421	6	AX332835
40	294	95.8	1421	9	HUMTCGXM
41	294	95.8	1655	9	BC039725 Homo sapi
42	292	95.1	757	9	HUMTCGCG
43	290	94.5	201	11	BV205008
44	290	94.5	300	9	HUMTCGCG
45	290	94.5	330	9	HUMTCGCG

#### ALIGNMENTS

RESULT 1  
HUMTCGCD1  
LOCUS  
DEFINITION Human T-cell receptor gamma-chain (TCRGC1) gene,  
C1-region, exon 1.  
ACCESSION M14996  
VERSION M14996.1 GI:339076  
C-region; T cell receptor gamma-chain; T-cell receptor; germline.  
SEGMENT 1 of 3  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 410)  
Lefranc,M.P., Forster,A. and Rabbits,T.H.  
Genetic polymorphism and exon changes of the constant regions of  
the human T-cell rearranging gene gamma  
Proc. Natl. Acad. Sci. U.S.A. 83 (24), 9596-9600 (1986)  
JOURNAL MEDLINE  
PUBMED 2879283  
COMMENT Original source text: Homo sapiens (clone: lambda-D19) DNA.  
Entry revisions for [1] kindly provided by M.-P. Lefranc,  
22-JUN-1988.

#### FEATURES

source  
1..410  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

/map="7p15-p14"  
/clone="lambda-D19"  
/cell\_line="ID-PLT"  
<41..370  
/gene="TCRCG1"  
/note="G00-120-408"  
/number=1

ORIGIN 205 bp upstream of BamHI site.

## Alignment Scores:

Pred. No.:	3.57e-30	Length:	410
Score:	307.00	Matches:	58
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-10-031-158-14 (1-58) x HUMTCGCD1 (1-410)

QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYsgInSerSer 20

DB 53 ATGCAGATGTTTCCCCCAAGCCACGATTTTCTTCTTCAATTGCTGAACAAAGCTCC 112

QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40

DB 113 AGAAGGCTGGAAACATACCTTGTCTTGTGAGAAATTTTCCCTGATGTTATTAAAGATAC 172

QY 41 IleGlyLysLysArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58

DB 173 ATTGGCAAGAAAGAAAGACCAACGATTTCTGGATCCAGAGGGGAAACACCA 226

RESULT 2

LOCUS AX884816 426 bp DNA linear PAT 18-DEC-2003

DEFINITION Sequence 679 from Patent EP1033401.

ACCESSION AX884816

VERSION AX884816.1 GI:40040355

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

Genet (FR)

Location/Qualifiers

1..426

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

267..>425

/note="unassigned protein product"

/codon\_start=1

/protein\_id="CAB98930.1"

/db\_xref="GI:40040356"

/translation="MKINDTYMKFSLVTPBKSIDKERRCIVRHNNKNGVDQETIIP

PKTDVITM"

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-10-031-158-14 (1-58) x AX884816 (1-426)

QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYsgInSerSer 20

DB 53 ATGCAGATGTTTCCCCCAAGCCACGATTTTCTTCTTCAATTGCTGAACAAAGCTCC 112

QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40

DB 113 AGAAGGCTGGAAACATACCTTGTCTTGTGAGAAATTTTCCCTGATGTTATTAAAGATAC 172

QY 41 IleGlyLysLysArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58

DB 173 ATTGGCAAGAAAGAAAGACCAACGATTTCTGGATCCAGAGGGGAAACACCA 226

RESULT 2

LOCUS AX884816 426 bp DNA linear PAT 18-DEC-2003

DEFINITION Sequence 679 from Patent EP1033401.

ACCESSION AX884816

VERSION AX884816.1 GI:40040355

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

Genet (FR)

Location/Qualifiers

1..426

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

267..>425

/note="unassigned protein product"

/codon\_start=1

/protein\_id="CAB98930.1"

/db\_xref="GI:40040356"

/translation="MKINDTYMKFSLVTPBKSIDKERRCIVRHNNKNGVDQETIIP

PKTDVITM"

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-10-031-158-14 (1-58) x AX884816 (1-426)

QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYsgInSerSer 20

DB 53 ATGCAGATGTTTCCCCCAAGCCACGATTTTCTTCTTCAATTGCTGAACAAAGCTCC 112

QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40

DB 113 AGAAGGCTGGAAACATACCTTGTCTTGTGAGAAATTTTCCCTGATGTTATTAAAGATAC 172

QY 41 IleGlyLysLysArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58

DB 173 ATTGGCAAGAAAGAAAGACCAACGATTTCTGGATCCAGAGGGGAAACACCA 226

RESULT 4

LOCUS CQ131517 470 bp DNA linear PAT 21-JAN-2004

DEFINITION Sequence 1539 from Patent WO0157276.

DB 94 ATGCAGATGTTTCCCCCAAGCCACGATTTTCTTCTTCAATTGCTGAACAAAGCTCC 153

QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40

DB 154 AGAAGGCTGGAAACATACCTTGTCTTGTGAGAAATTTTCCCTGATGTTATTAAAGATAC 213

QY 41 IleGlyLysLysArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58

DB 214 ATTGGCAAGAAAGAAAGACCAACGATTTCTGGATCCAGAGGGGAAACACCA 267

## RESULT 3

BD024426 426 bp DNA linear PAT 27-AUG-2002

LOCUS BD024426

DEFINITION Sequence tag and encoded human protein.

ACCESSION BD024426

VERSION BD024426.1 GI:22565649

KEYWORDS JP 2001269182-A/672

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Edwards, J.B.D.M., Duclair, E. and Jordan, J.Y.

AUTHORS 1 (bases 1 to 426)

TITLE Sequence tag and encoded human protein

JOURNAL Patent: JP 2001269182-A 672 02-OCT-2001.

COMMENT

OS Homo sapiens (human)

PN JP 2001269182-A/672

PD 02-OCT-2001

PF 24-FEB-2000 JP 2000118773

PI 26-FEB-1999 US 60/122487

PI JEAN BAPTIST DUMAS MILNE EDWARDS, EIMERIC DUCLAIR, JEAN YVES

PI JORDAN

PC C12N15/09, C07K14/435, C07K16/18, C12N1/15, C12N1/19, C12N1/21, PC

C12N5/10.

PC C12P21/02, C12P21/08, C12Q1/68//G06F17/30, C12N15/00, C12N5/00, PC

G06F15/40

CC

FH Key

FT CDS

Location/Qualifiers

1..426

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

## FEATURES

source

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-10-031-158-14 (1-58) x BD024426 (1-426)

QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYsgInSerSer 20

DB 94 ATGCAGATGTTTCCCCCAAGCCACGATTTTCTTCTTCAATTGCTGAACAAAGCTCC 153

QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40

DB 154 AGAAGGCTGGAAACATACCTTGTCTTGTGAGAAATTTTCCCTGATGTTATTAAAGATAC 213

QY 41 IleGlyLysLysArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58

DB 214 ATTGGCAAGAAAGAAAGACCAACGATTTCTGGATCCAGAGGGGAAACACCA 267

ACCESSION CQ131517  
VERSION CQ131517.1 GI:41088873  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1  
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.  
TITLE Human genome-derived single exon nucleic acid probes useful for  
analysis of gene expression in human bone marrow  
JOURNAL Patent: WO 0157276-A 1539 09-AUG-2001;  
Acemica, Inc. (US)

FEATURES  
source location/Qualifiers  
1..470  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="MAP TO AF159056.1-EXPRESSED IN BONE MARROW, SIGNAL  
= 2.2"

ORIGIN  
Alignment Scores:  
Pred. No.: 4.16e-30 Length: 470  
Score: 307.00 Matches: 58  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0

US-10-031-158-14 (1-58) x CQ131517 (1-470)

Qy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20  
Db 203 ATGCAGATGTTCCGCCAAGCCCACTATTCTTCTTCCTGAAACAAAGCTCC 262

Qy 21 ArgArgLeuGlnIsthrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40  
Db 263 AGAAGGCTGGAAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAAGATAC 322

Qy 41 IleGlyLysLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58  
Db 323 ATTGGCAAGAAAGAAAGAGCAACGATTCGGATCCAGAGGGGAAACACCA 376

RESULT 5  
LOCUS CQ290455 470 bp DNA linear PAT 23-JAN-2004  
DEFINITION Sequence 1560 from Patent WO0186003.  
ACCESSION CQ290455  
VERSION CQ290455.1 GI:41251032  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1  
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.  
TITLE Human genome-derived single exon nucleic acid probes useful for  
analysis of gene expression in human lung  
JOURNAL Patent: WO 0186003-A 1560 15-NOV-2001;  
Acemica, Inc. (US)

FEATURES  
source location/Qualifiers  
1..470  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="MAP TO AF159056.1-EXPRESSED IN LUNG, SIGNAL = 1.1"

ORIGIN  
Alignment Scores:  
Pred. No.: 4.16e-30 Length: 470  
Score: 307.00 Matches: 58  
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0

US-10-031-158-14 (1-58) x CQ290455 (1-470)

Qy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20  
Db 203 ATGCAGATGTTCCGCCAAGCCCACTATTCTTCTTCCTGAAACAAAGCTCC 262

Qy 21 ArgArgLeuGlnIsthrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40  
Db 263 AGAAGGCTGGAAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAAGATAC 322

Qy 41 IleGlyLysLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58  
Db 323 ATTGGCAAGAAAGAAAGAGCAACGATTCGGATCCAGAGGGGAAACACCA 376

RESULT 6  
LOCUS CQ132293 477 bp DNA linear PAT 21-JAN-2004  
DEFINITION Sequence 2315 from Patent WO0157276.  
ACCESSION CQ132293  
VERSION CQ132293.1 GI:41089649  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1  
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.  
TITLE Human genome-derived single exon nucleic acid probes useful for  
analysis of gene expression in human bone marrow  
JOURNAL Patent: WO 0157276-A 2315 09-AUG-2001;  
Acemica, Inc. (US)

FEATURES  
source location/Qualifiers  
1..477  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="MAP TO AF159056.1-EXPRESSED IN BONE MARROW, SIGNAL  
= 1.4"

ORIGIN  
Alignment Scores:  
Pred. No.: 4.23e-30 Length: 477  
Score: 307.00 Matches: 58  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0

US-10-031-158-14 (1-58) x CQ132293 (1-477)

Qy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20  
Db 212 ATGCAGATGTTCCGCCAAGCCCACTATTCTTCTTCCTGAAACAAAGCTCC 271

Qy 21 ArgArgLeuGlnIsthrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40  
Db 272 AGAAGGCTGGAAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAAGATAC 331

Qy 41 IleGlyLysLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58  
Db 332 ATTGGCAAGAAAGAAAGAGCAACGATTCGGATCCAGAGGGGAAACACCA 385

RESULT 7  
LOCUS CQ200006 477 bp DNA linear PAT 21-JAN-2004  
DEFINITION Sequence 2283 from Patent WO0157271.  
ACCESSION CQ200006  
VERSION CQ200006.1 GI:41185670  
KEYWORDS

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1  
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.  
TITLE Human genome-derived single exon nucleic acid probes useful for  
analysis of gene expression in human breast and bc 474 cells  
JOURNAL Patent: WO 0157271-A 2283 09-AUG-2001;  
Aeomica, Inc. (US)

FEATURES  
source location/Qualifiers  
1..477  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="MAP TO AF159056.1-EXPRESSED IN BT474, SIGNAL = 1.1"

ORIGIN

Alignment Scores:  
Pred. No.: 4,23e-30 Length: 477  
Score: 307.00 Matches: 58  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0

US-10-031-158-14 (1-58) x CQ200006 (1-477)

Qy 1 MetGlnMetPheProPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20  
Db 212 ATGCAGATGTTTCCCCCAAGCCCACTATTTCTTCCTCAATTCCTGAAACAAAGCTCC 271

Qy 21 ArgArgLeuGlnIuHsIthRphValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40  
Db 272 AGAAGGCTGGAACATACCTTGTCTTCTTGAAATTTTCCCTGATGTATTAAGATAC 331

Qy 41 IlegIyLysLysArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58  
Db 332 ATTGGCAAGAAAGAGACACGATTCCTGGATCCAGAGAGGGAACACCA 385

RESULT 8  
CQ291096 477 bp DNA linear PAT 23-JAN-2004  
DEFINITION Sequence 2201 from Patent WO0186003.  
ACCESSION CQ291096  
VERSION CQ291096.1 GI:41251673  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1  
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.  
TITLE Human genome-derived single exon nucleic acid probes useful for  
analysis of gene expression in human lung  
JOURNAL Patent: WO 0186003-A 2201 15-NOV-2001;  
Aeomica, Inc. (US)

FEATURES  
source location/Qualifiers  
1..477  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="MAP TO AF159056.1-EXPRESSED IN LUNG, SIGNAL = 5.3"

ORIGIN

Alignment Scores:  
Pred. No.: 4,23e-30 Length: 477  
Score: 307.00 Matches: 58  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0

US-10-031-158-14 (1-58) x CQ291096 (1-477)

Qy 1 MetGlnMetPheProPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20  
Db 212 ATGCAGATGTTTCCCCCAAGCCCACTATTTCTTCCTCAATTCCTGAAACAAAGCTCC 271

Qy 21 ArgArgLeuGlnIuHsIthRphValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40  
Db 272 AGAAGGCTGGAACATACCTTGTCTTCTTGAAATTTTCCCTGATGTATTAAGATAC 331

Qy 41 IlegIyLysLysArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58  
Db 332 ATTGGCAAGAAAGAGACACGATTCCTGGATCCAGAGAGGGAACACCA 385

RESULT 9  
CQ490917/c 539 bp DNA linear PAT 30-JAN-2004  
LOCUS CQ490917  
DEFINITION Sequence 22784 from Patent WO0160860.  
ACCESSION CQ490917  
VERSION CQ490917.1 GI:41456536  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1  
AUTHORS Schlegel, R., Endege, W.O. and Monahan, J.E.  
TITLE Genes differentially expressed in human prostate cancer and their  
use  
JOURNAL Patent: WO 0160860-A 22784 23-AUG-2001;  
Millennium Predictive Medicine, Inc. (US)

FEATURES  
source location/Qualifiers  
1..539  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN

Alignment Scores:  
Pred. No.: 4,84e-30 Length: 539  
Score: 307.00 Matches: 58  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0

US-10-031-158-14 (1-58) x CQ490917 (1-539)

Qy 1 MetGlnMetPheProPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20  
Db 410 ATGCAGATGTTTCCCCCAAGCCCACTATTTCTTCCTCAATTCCTGAAACAAAGCTCC 351

Qy 21 ArgArgLeuGlnIuHsIthRphValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40  
Db 350 AGAAGGCTGGAACATACCTTGTCTTCTTGAAATTTTCCCTGATGTATTAAGATAC 291

Qy 41 IlegIyLysLysArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58  
Db 290 ATTGGCAAGAAAGAGACACGATTCCTGGATCCAGAGAGGGAACACCA 237

RESULT 10  
CQ496769/c 539 bp DNA linear PAT 30-JAN-2004  
LOCUS CQ496769/c  
DEFINITION Sequence 28636 from Patent WO0160860.  
ACCESSION CQ496769  
VERSION CQ496769.1 GI:41462398  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1  
AUTHORS Schlegel, R., Endege, W.O. and Monahan, J.E.

TITLE Genes differentially expressed in human prostate cancer and their use

JOURNAL Patent: WO 0160860-A 28636 23-AUG-2001;  
Millennium Predictive Medicine, Inc. (US)

FEATURES  
source  
1. .539  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN

Alignment Scores:  
Pred. No.: 4,84e-30 Length: 539  
Score: 307.00 Matches: 58  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0  
DB: 6

US-10-031-158-14 (1-58) x CQ496769 (1-539)

Qy 1 MetGlnMetPheProPheSerProLeuPhePheLeuGlnLeuLeuLysGlnSerSer 20  
|||  
Db 410 ATGCAGATGTTTCCCCCAAGCCCACTATTTTCTTCTTCAATTCGTAACAAAGCTCC 351  
|||

Qy 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40  
|||  
Db 350 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAGATAC 291  
|||

Qy 41 IleGlyLysLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58  
|||  
Db 290 ATTGGCAAGAAAGAGACACAGATTCGTGGATCCAGAGGGGACACCA 237  
|||

RESULT 11

AJ583012 723 bp mRNA linear PRI 01-OCT-2003  
LOCUS Homo sapiens partial mRNA for T-cell receptor gamma chain (TCRG  
DEFINITION gene). clone 15.1.21.Bob.  
ACCESSION AJ583012.1 GI:37495406  
VERSION T-cell receptor gamma chain; TCRG gene.  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1  
Green,A.E., Lissina,A., Hutchinson,S.L., Temple,B., Boulter,J.M.,  
Price,D.A. and Sewell,A.K.  
Recognition of non-peptide antigens by human Vgamma9 Vdelta2 TCRs  
requires contact with a cell of human origin  
Unpublished  
2 (bases 1 to 723)

JOURNAL  
TITLE Direct Submission  
AUTHORS Green,A.E.  
JOURNAL Submitted (16-SEP-2003) Green A.E., Nuffield Department of  
Medicine, University of Oxford, Peter Medawar Building, South Parks  
Rd, Oxford, OX1 3SY, UNITED KINGDOM  
location/Qualifiers  
1. .723  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="15.1.21.Bob"  
/cell\_type="gamma/delta T cell"  
/rearranged  
1. .723  
/gene="TCRG"  
1. .>723  
/gene="TCRG"  
/codon\_start=1  
/evidence=experimental  
/product="T-cell receptor gamma chain"  
/protein\_id="CAE47524.1"

FEATURES  
source  
1. .723  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="15.1.21.Bob"  
/cell\_type="gamma/delta T cell"  
/rearranged  
1. .723  
/gene="TCRG"  
1. .>723  
/gene="TCRG"  
/codon\_start=1  
/evidence=experimental  
/product="T-cell receptor gamma chain"  
/protein\_id="CAE47524.1"

/db\_xref="GI:37495407"

/translation="MAGHLEQPOISSTKTLSTKARLECVSGITISATSYYWYRRPG  
EVIOPLVISYSDGVTRKESGISGKPEVRIPESTSLTHVBEKODIATVYCALBE  
AKGKIKRVPSTKILITDKDADAVSPRTPLPSIATTKOKGTYCLIEKPP  
DVTKIHWQEKSNFTILGSGQGNMTKNDITMKSFWLTPESLDKEHRCIVRHNNKN  
GVDOEIIFFPIKTVITMPKDN"

4. .306  
/gene="TCRG"  
/note="TRGV9"  
313. .363  
/gene="TCRG"  
/note="TRGJP"  
364. .>723  
/gene="TCRG"  
/note="TRGCI"

ORIGIN

Alignment Scores:  
Pred. No.: 6.72e-30 Length: 723  
Score: 307.00 Matches: 58  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0  
DB: 9

US-10-031-158-14 (1-58) x AJ583012 (1-723)

Qy 1 MetGlnMetPheProPheSerProLeuPhePheLeuGlnLeuLeuLysGlnSerSer 20  
|||  
Db 377 ATGCAGATGTTTCCCCCAAGCCCACTATTTTCTTCTTCAATTCGTAACAAAGCTCC 436  
|||

Qy 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40  
|||  
Db 437 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAGATAC 496  
|||

Qy 41 IleGlyLysLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58  
|||  
Db 497 ATTGGCAAGAAAGAGACACAGATTCGTGGATCCAGAGGGGACACCA 550  
|||

RESULT 12

AJ583014 726 bp mRNA linear PRI 01-OCT-2003  
LOCUS Homo sapiens partial mRNA for T-cell receptor gamma chain (TCRG  
DEFINITION gene). clone 16.1.2.Wendy.  
ACCESSION AJ583014.1 GI:37495410  
VERSION T-cell receptor gamma chain; TCRG gene.  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1  
Green,A.E., Lissina,A., Hutchinson,S.L., Temple,B., Boulter,J.M.,  
Price,D.A. and Sewell,A.K.  
Recognition of non-peptide antigens by human Vgamma9 Vdelta2 TCRs  
requires contact with a cell of human origin  
Unpublished  
2 (bases 1 to 726)

JOURNAL  
TITLE Direct Submission  
AUTHORS Green,A.E.  
JOURNAL Submitted (16-SEP-2003) Green A.E., Nuffield Department of  
Medicine, University of Oxford, Peter Medawar Building, South Parks  
Rd, Oxford, OX1 3SY, UNITED KINGDOM  
location/Qualifiers  
1. .726  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="16.1.2.Wendy"  
/cell\_type="gamma/delta T cell"  
/rearranged  
1. .726  
/gene="TCRG"

FEATURES  
source  
1. .726  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="16.1.2.Wendy"  
/cell\_type="gamma/delta T cell"  
/rearranged  
1. .726  
/gene="TCRG"

CDS

1. .>726  
/gene="TCRG"  
/codon\_start=1  
/product="T-cell receptor gamma chain"  
/protein\_id="CA847526.1"  
/db\_xref="GI:37495411"  
/translation="MAGHLKQPOISSTYKLSKTLARLCEVSGITISATSVYWRERPG  
EVIGFLVSIYDGVTRKESGIPSGFEVDRIPESTSTLTITHNEKODIATVYCALME  
VTBLGKIKIVFGPGTKLIITDKLDADVSPPRTIFLPSIATKLOKAGTYCLKEFP  
PDVIRKHQEKSNITIGSOGNTMKNTDVTMKRSWLVPESLDKEHRCIVREHNK  
NGVDDELFPPIKIDVITMDPKDN"  
4. .309  
/gene="TCRG"  
/note="TRGV9"  
313. .366  
/gene="TCRG"  
/note="TRGJP"  
367. .>726  
/gene="TCRG"  
/note="TRGCI"

C\_region

ORIGIN

Alignment Scores:

Pred. No.: 6.79e-30 Length: 726  
Score: 307.00 Matches: 58  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0

US-10-031-158-14 (1-58) x AJ583014 (1-726)

Qy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuLysGlnSerSer 20  
Db 380 ATGCAGAGTGTTCCTCCCAAGCCCACTATTTCTTCCTCAATCTGAAACAAAGCTCC 439

Qy 21 ArgArgLeuGlnLhiThrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40  
Db 440 AGAAGGCTGGAAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAGATAC 499

Qy 41 IlleGlyValLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58  
Db 500 ATTGGCAAGAAAGAAAGAGCAACAGCATTTCTGGATCCAGAGGAGAACCA 553

RESULT 13

189883 189883 825 bp DNA linear PAT 10-AUG-1998

LOCUS DEFINITION Sequence 18 from patent US 5723309.

ACCESSION 189883

VERSION 189883.1 GI:3409823

KEYWORDS

SOURCE Unknown.

ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 825)

AUTHORS Bonneville,M.

TITLE Production of subunits of soluble T cell receptors by co-transfection

JOURNAL Patent: US 5723309-A 18-03-MAR-1998;

FEATURES

source Location/Qualifiers

1. .825  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Alignment Scores:

Pred. No.: 7.79e-30 Length: 825  
Score: 307.00 Matches: 58  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0

US-10-031-158-14 (1-58) x 189883 (1-825)

Qy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuLysGlnSerSer 20  
Db 440 ATGCAGAGTGTTCCTCCCAAGCCCACTATTTCTTCCTCAATCTGAAACAAAGCTCC 499

Qy 21 ArgArgLeuGlnLhiThrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40  
Db 500 AGAAGGCTGGAAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAGATAC 559

Qy 41 IlleGlyValLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58  
Db 560 ATTGGCAAGAAAGAAAGAGCAACAGATTTCTGGATCCAGAGGAGAACCA 613

RESULT 14

HSTRGSA 825 bp mRNA linear PRI 09-SEP-1993

LOCUS DEFINITION H.sapiens mRNA for soluble gamma TCR.

ACCESSION X72500.1 GI:298106

VERSION X72500.1

KEYWORDS soluble receptor; T cell receptor gamma chain.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 825)

AUTHORS Davodeau,F., Houde,I., Boulot,G., Romagne,F., Necker,A., Canavo,N., Peyrat,M.A., Haller,M.M., Vie,H., Jacques,Y., Martuza,R. and Bonneville,M.

TITLE Secretion of disulfide-linked human T-cell receptor gamma delta heterodimers

JOURNAL J. Biol. Chem. 268 (21), 15455-15460 (1993)

MEDLINE 93340140

PUBMED 8340374

REFERENCE 2 (bases 1 to 825)

AUTHORS Davodeau,F.

TITLE Direct Submission

JOURNAL Submitted (28-APR-1993) F. Davodeau, INSERM U211, Institut de Biologie, 9 Quai Moncousu, 44035 Nantes Cedex 01, FRANCE

FEATURES

source Location/Qualifiers

1. .825  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="7"  
/map="7q15"  
/cell\_type="T lymphocyte"  
/tissue\_type="blood"  
1. .825  
/codon\_start=1  
/product="gamma-delta T-cell receptor"  
/protein\_id="CA51165.1"  
/db\_xref="GI:298107"  
/translation="MELSLHASTLVAVGALCVGAGHLEQPOISSTYKLSKTLARLCEV  
VSGITISATSVYWRERPGEVIGFLVSIYDGVTRKESGIPSGFEVDRIPESTSTLTITHNEKODIATVYCALME  
VTBLGKIKIVFGPGTKLIITDKLDADVSPPRTIFLPSIATKLOKAGTYCLKEFPDVKIRKHQEKSNITIGSOGNTMKNTDVTMKRSWL  
VPEKSLDKERHCIVREHNKGVQEIFPPIKIDVITMDPKDNCKSDANTLLIQ"

V\_region 1. .364  
/product="V-gamma-9"

N\_region 365. .368  
/product="V-gamma-9"

J\_region 369. .426  
/product="J-gamma-P"

C\_region 427. .825  
/product="C-gamma"

ORIGIN

Alignment Scores:

Pred. No.: 7.79e-30 Length: 825  
Score: 307.00 Matches: 58  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0

DB: 9 Gaps: 0

US-10-031-158-14 (1-58) x HSTCRGSA (1-825)

Qy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuLYGlnSerSer 20  
|||||  
Db 440 ATGCAGATGTTTCCCGCAAGCCCACTATTTTCTTCTTCATTCGTAACCAAGCTCC 499

Qy 21 ArgArgLeuGlnIuHstItrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40  
|||||  
Db 500 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAAGATAC 559

Qy 41 IlleGlyValSerArgArgAlaThrArgPheTrpAspProArgArgIYThrPro 58  
|||||  
Db 560 ATTGGAGAGAAAGAGACACGATTTCTGGATTCACGAGGAGGACACCA 613

RESULT 15  
LOCUS HUMTCRGAD 958 bp mRNA linear PRI 14-JAN-1995  
DEFINITION Homo sapiens (clone HGP02) T cell receptor gamma-chain mRNA, Cl region.  
ACCESSION M27334  
VERSION M27334.1 GI:540463  
KEYWORDS T cell receptor gamma-chain; constant region.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
Yoshikai.Y., Toyonaga.B., Koga.Y., Kimura.N., Griesser.H. and Mak.T.W.  
REPERIOICE of the human T cell gamma genes: high frequency of nonfunctional transcripts in thymus and mature T cells  
Eur. J. Immunol. 17 (1), 119-126 (1987)  
JOURNAL MEDLINE 87133835  
PUBMED 2949984  
COMMENT On Sep 21, 1994 this sequence version replaced gi:341672.  
Original source text: Homo sapiens CDNA to mRNA.  
FEATURES  
source  
1..958  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/map="7p15-p14"  
/clone="HGP02"  
/cell\_type="PHA-stimulated peripheral T cell"  
1..958  
/gene="TCRCG1"  
<1..522  
/gene="TCRCG1"  
/codon\_start=1  
/product="T cell receptor gamma chain"  
/protein\_id="AA6114.1"  
/db\_xref="GI:540464"  
/translation="DKQLDADVSRKPTIFLPSIAETKIQKAGTYLCLEKFPVYIKI HMEKSKNTLIGSQENMTKNTDYMKFSMLTVEKSLDKRHCIVRHNNKGVDSIIP IIPPIKTDVITMDPKNCSKDANDTLLQLTNTSAYYVYLLLLKSVYFAITTCCL LRTAFCCNGEKS"  
941..946  
/gene="TCRCG1"  
/note="G00-120-408"  
958  
/gene="TCRCG1"  
/note="G00-120-408"  
polya\_signal  
941..946  
/gene="TCRCG1"  
/note="G00-120-408"  
polya\_site  
958  
/gene="TCRCG1"  
/note="G00-120-408"  
ORIGIN  
Alignment Scores:  
Pred. No.: 9.21e-30 Length: 958  
Score: 307.00 Matches: 58  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0

DB: 9 Gaps: 0

US-10-031-158-14 (1-58) x HUMTCRGAD (1-958)

Qy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuLYGlnSerSer 20  
|||||  
Db 14 ATGCAGATGTTTCCCGCAAGCCCACTATTTTCTTCTTCATTCGTAACCAAGCTCC 73

Qy 21 ArgArgLeuGlnIuHstItrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40  
|||||  
Db 74 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAAGATAC 133

Qy 41 IlleGlyValSerArgArgAlaThrArgPheTrpAspProArgArgIYThrPro 58  
|||||  
Db 134 ATTGGAGAGAAAGAGACACGATTTCTGGATTCACGAGGAGGACACCA 187

RESULT 16  
LOCUS AX074415 1027 bp DNA linear PAT 06-FEB-2001  
DEFINITION Sequence 13 from Patent WO0104309.  
ACCESSION AX074415  
VERSION AX074415.1 GI:12710553  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
Pastan,I., Essand,M., Lee,B., Vasmatazis,G. and Wolfgang,C.  
T-cell receptor \_g(g) alternate reading frame protein, (tarp) and uses thereof  
Patent: WO 0104309-A 13 18-JAN-2001;  
JOURNAL UNITED STATES GOVERNMENT (US)  
FEATURES  
source  
1..1027  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
74..250  
/note="unnamed protein product; Coding region for PS-TCR gamma 1 polypeptide (TRAP)"  
/codon\_start=1  
/protein\_id="CAC28474.1"  
/db\_xref="GI:12710554"  
/translation="MQMPPSPPLFFLIQLKQSRRLHFTVFLRNSIMLRITGKK RRAIRFWDPRGRTP"  
247..582  
/note="unnamed protein product; Coding region for PS-TCR gamma 2 polypeptide (deduced amino acid sequence not displayed along with DNA sequence, due to overlapping CDS's)"  
/codon\_start=1  
/protein\_id="CAC28474.1"  
/db\_xref="GI:12710555"  
/translation="MKTNDYMKRSMULTVEKSLDKRHCIVRHNNKGVDSIIP PIRTDVITMDPKNCSKDANDTLLQLTNTSAYYVYLLLLKSVYFAITTCCLRT AFCCNGEKS"  
ORIGIN  
Alignment Scores:  
Pred. No.: 9.95e-30 Length: 1027  
Score: 307.00 Matches: 58  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-031-158-14 (1-58) x AX074415 (1-1027)

Qy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuLYGlnSerSer 20  
|||||  
Db 74 ATGCAGATGTTTCCCGCAAGCCCACTATTTTCTTCTTCATTCGTAACCAAGCTCC 133

Qy 21 ArgArgLeuGluHstHrPheValPheLeuArgAnPheSerLeuMetLeuArgTyr 40  
 Db 134 AGAAGGCTGGAACATACCTTGTCTTGTGAATAATTTTCCGTGATGTATTAAGATAC 193  
 Qy 41 lIleGlyLysLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58  
 Db 194 ATTGGCAAGAAAAGAGAGCAACGATTTGGGATCCAGAGGGGAACACCA 247

RESULT 17  
 AF151103 1027 bp mRNA linear PRI 03-NOV-2000  
 LOCUS Homo sapiens TCRgamma alternate reading frame protein (TCRG) mRNA,  
 complete cds.  
 ACCESSION AF151103  
 VERSION AF151103.1 GI:5758136  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 Esasand,M., Vasmatazle,G., Brinkmann,U., Duray,P., Lee,B. and  
 Pastan,I.  
 TITLE High expression of a specific T-cell receptor gamma transcript in  
 epithelial cells of the prostate  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (16), 9287-9292 (1999)  
 MEDLINE 99362750  
 PUBMED 10430935  
 REFERENCE 2 (bases 1 to 1027)  
 AUTHORS Wolfgang,C.D., Esasand,M., Vincent,J.J., Lee,B. and Pastan,I.  
 TITLE TARP: a nuclear protein expressed in prostate and breast cancer  
 cells derived from an alternate reading frame of the T cell  
 receptor gamma chain locus  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (17), 9437-9442 (2000)  
 MEDLINE 20402552  
 PUBMED 10931945  
 REFERENCE 3 (bases 1 to 1027)  
 AUTHORS Esasand,M., Vasmatazle,G., Brinkmann,U., Duray,P., Lee,B. and  
 Pastan,I.  
 TITLE Direct Submission  
 JOURNAL Submitted (13-MAY-1999) Division of Basic Sciences, Laboratory of  
 Molecular Biology, National Institutes of Health, National Cancer  
 Institute, 37 Convent Drive, Bethesda, MD 20892, USA  
 FEATURES  
 source  
 1..1027  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /note="transcript from an unrearranged TCRG locus"  
 gene  
 1..1027  
 /gene="TCRG"  
 J\_segment  
 1..60  
 /gene="TCRG"  
 C\_region  
 61..390  
 /note="gamma 1.2"  
 /gene="TCRG"  
 CDS  
 74..250  
 /note="gamma 1; corresponds to exon I"  
 /gene="TCRG"  
 /note="TARP"  
 /codon\_start=1  
 /product="TCRgamma alternate reading frame protein"  
 /protein\_id="AAG29337.1"  
 /db\_xref="GI:11093514"  
 /translation="MQMPPPSLFFFLQLKOSRRLHPTVFLNFSMLRLRYGKK  
 RRATFMDPRRGT"  
 C\_region  
 391..438  
 /gene="TCRG"  
 /note="gamma 1; corresponds to exon II"  
 C\_region  
 439..579  
 /gene="TCRG"  
 /note="gamma 1; corresponds to exon III"

Alignment Scores:  
 Pred. No.: 9.95e-10 Length: 1027  
 Score: 307.00 Matches: 58  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: Gaps: 0

US-10-031-158-14 (1-58) x AF151103 (1-1027)

Qy 1 MetGlnMetPheProProSerProLeuPhePheLeuGlnLeuLysGlnSerSer 20  
 Db 74 ATGCAGATGTTCCCGCAAGCCACTATTTCTTCCTTCATTCGTAACCAAGCTCC 133

Qy 21 ArgArgLeuGluHstHrPheValPheLeuArgAnPheSerLeuMetLeuArgTyr 40  
 Db 134 AGAAGGCTGGAACATACCTTGTCTTGTGAATAATTTTCCGTGATGTATTAAGATAC 193

Qy 41 lIleGlyLysLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58  
 Db 194 ATTGGCAAGAAAAGAGAGCAACGATTTGGGATCCAGAGGGGAACACCA 247

RESULT 18  
 BC062761 1054 bp mRNA linear PRI 16-DEC-2003  
 LOCUS Homo sapiens T cell receptor gamma variable 9, mRNA (cDNA clone  
 IMAGE:4248993), partial cds.  
 ACCESSION BC062761  
 VERSION BC062761.1 GI:38540986  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 Klausner,R.D., Collins,F.S., Wagner,L., Shanmen,C.M., Schuler,G.D.,  
 Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhac,N.K.,  
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,  
 Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,  
 Stapleton,M., Soares,M.B., Bonaldo,M.F., Cabavant,T.L.,  
 Schaefer,T.E., Brownstein,M.J., Uedlin,T.B., Tohilyuk,I.S.,  
 Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,  
 Abramson,R.D., Mullahy,S.J., Bosak,S.A., McGowan,P.J.,  
 McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,  
 Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Bulky,S.W.,  
 Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,  
 Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,  
 Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,  
 Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,  
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,  
 Butterfield,Y.S., Krzywicki,M.I., Skalka,J., Smalls,D.E.,  
 Schnerch,A., Schein,J.B., Jones,S.J. and Marra,M.A.  
 TITLE Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 PUBMED 12477932  
 REFERENCE 2 (bases 1 to 1054)  
 AUTHORS Strausberg,R.  
 TITLE Direct Submission  
 JOURNAL Submitted (25-NOV-2003) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: http://mgc.nci.nih.gov  
 Contact: MGC help desk  
 Email: c9apbs-rc@mail.nih.gov  
 Tissue Procurement: CLOMTECH  
 cDNA Library Preparation: CLOMTECH Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Genome Sequence Centre,  
 BC Cancer Agency, Vancouver, BC, Canada

info@dcgsc.bc.ca  
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,  
Susanna Chen, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,  
Leticia Hsiao, Martin Krzywinski, Keta Kutsche, Oliver Lee, Soo  
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven  
Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline  
Schein, Duane Smalls, Michael Smith, Lorraine Spence, Jeff Scott,  
Michael Thorne, Miranada Tsai, Natasha van den Bosch, Jill Vardy,  
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LINL ac: <http://image.lnl.gov>  
Series: IRAL Plate: 51 Row: e Column: 16  
This clone was selected for full length sequencing because it  
passed the following selection criteria. Hexamer frequency ORF  
analysis.

#### FEATURES

source  
1. .1054  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_image="424893"  
/issue\_type="Prostate"  
/clone\_id="NH MGC\_83"  
/lab\_host="DH10B"  
/note="Vector: pDNR-LIB"  
1. .1054  
/gene="TRGV9"  
/note="synonyms: TCRGV9, V2"  
/db\_xref="IMG/IMG:TRGV9"  
/db\_xref="LOCUSID:6983"  
1. .577  
/gene="TRGV9"  
/codon\_start=2  
/product="TRGV9 protein"  
/protein\_id="AA62761.1"  
/db\_xref="GI:38540987"  
/db\_xref="IMG/IMG:TRGV9"  
/db\_xref="LOCUSID:6983"  
/translation="GGQKIKVFGPGTKLIITDKOLDADVSFKPTLPSTIAETKLOK  
AGTYICLKEFPPIVIKIMQEKSNILGSEBNTKNTYKFKFVLTVPKSLDK  
EHRCIVHNNKGVDEIIPPIKTDVITVDPPDNCSKDNDTLILQLTWTSAYMY  
LILLLKSVYFAIITTCILRTATPCNGEKS"  
74. .364  
/gene="TRGV9"  
/note="IGC: Region: Immunoglobulin domain constant region  
subfamily"  
/db\_xref="CDD:cd00098"

#### misc\_feature

/gene="TRGV9"  
/note="IGC: Region: Immunoglobulin domain constant region  
subfamily"  
/db\_xref="CDD:cd00098"

#### ORIGIN

Alignment Scores:  
Pred. No.: 1.02e-29 Length: 1054  
Score: 307.00 Matches: 58  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-10-031-158-14 (1-58) x BC062761 (1-1054)

OY 1 MetGlnMetPheProPheSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20  
DB 69 ATGAGAGATTTCCTCCCAAGCCCACTATTTCTTCTTCAATGCTGAAACAAGCTCC 128  
OY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40  
DB 129 AGAAGGCTGGAAACATACCTTGTCTTCTGAGAAATTTTCCCTGATGTTATTAAGATAC 188  
OY 41 TleGlyLyLySArGArGAlaThArGpHeTrPaSPProArGArGlyThPro 58  
DB 189 ATTGGCAAGAAAGAAAGACAAACGATTCTGGATCCAGAGGGGGAACACCA 242  
RESULT 19

CQ493215/c  
LOCUS CQ493215 1316 bp DNA linear PAT 30-JAN-2004  
DEFINITION Sequence 25082 from Patent WO0160860.  
ACCESSION CQ493215  
VERSION CQ493215.1 GI:41458834  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

#### REFERENCE

1 Schlegel, R., Endege, W.O. and Monahan, J.E.  
Genes differentially expressed in human prostate cancer and their  
use

JOURNML Patent: WO 0160860-A 25082 23-AUG-2001;  
Millennium Predictive Medicine, Inc. (US)

#### FEATURES

source  
1. .1316  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

#### ORIGIN

Alignment Scores:  
Pred. No.: 1.31e-29 Length: 1316  
Score: 307.00 Matches: 58  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-031-158-14 (1-58) x CQ493215 (1-1316)

OY 1 MetGlnMetPheProPheSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20  
DB 562 ATGAGAGATTTCCTCCCAAGCCCACTATTTCTTCTTCAATGCTGAAACAAGCTCC 503  
OY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40  
DB 502 AGAAGGCTGGAAACATACCTTGTCTTCTGAGAAATTTTCCCTGATGTTATTAAGATAC 443  
OY 41 TleGlyLyLySArGArGAlaThArGpHeTrPaSPProArGArGlyThPro 58  
DB 442 ATTGGCAAGAAAGAAAGACAAACGATTCTGGATCCAGAGGGGGAACACCA 389

#### RESULT 20

CQ494607/c CQ494607 1316 bp DNA linear PAT 30-JAN-2004  
LOCUS CQ494607 Sequence 26474 from Patent WO0160860.  
DEFINITION CQ494607  
ACCESSION CQ494607  
VERSION CQ494607.1 GI:41460226  
KEYWORDS

SOURCE  
ORGANISM  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

#### REFERENCE

1 Schlegel, R., Endege, W.O. and Monahan, J.E.  
Genes differentially expressed in human prostate cancer and their  
use

JOURNML Patent: WO 0160860-A 26474 23-AUG-2001;  
Millennium Predictive Medicine, Inc. (US)

#### FEATURES

source  
1. .1316  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

#### ORIGIN

Alignment Scores:  
Pred. No.: 1.31e-29 Length: 1316  
Score: 307.00 Matches: 58  
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-10-031-158-14 (1-58) x CG494607 (1-1316)

Qy 1 MetGlnMetPheProPheSerProLeuPhePheLeuGlnLeuLeuLyGlnSerSer 20  
 Db 562 ATGCAGATGTTCCCGCCAGCCCACTATTTTCTTCCTCAATGCTGAAACCAAGCTCC 503

Qy 21 ArgArgLeuGlnHisThrPheValPheLeuArgAenPheSerLeuMetLeuLeuArgTyr 40  
 Db 502 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTATTAAGATAC 443

Qy 41 lIleGlyLysLysArgArgAlaThrArgPheTrpApeProArgArgGlyThrPro 58  
 Db 442 ATTGGCAAGAAAGAAAGCAACGATTCGTGGATTCAGAGGAGGAAACACCA 389

RESULT 21  
 BC072387 1470 bp mRNA linear PRI 30-JUN-2004  
 LOCUS IMAGE:5587705, partial cds.  
 DEFINITION Homo sapiens T cell receptor gamma variable 9, mRNA (cDNA clone  
 IMAGE:5587705), partial cds.  
 ACCESSION BC072387  
 VERSION BC072387.1 GI:47938917  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 1470)  
 Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,  
 Klausner,R.D., Collins,F.S., Wagner,K.H., Shenmen,C.M., Schuler,G.D.,  
 Altschul,S.F., Zeeberg,B., Buetow,K.H., Scheffer,C.F., Bhat,N.K.,  
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hejeh,F.,  
 Diachenko,L., Marusina,K., Farmer,A.A., Rubin,J., Hong,L.,  
 Slepietson,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,  
 Schenier,T.E., Brownstein,M.J., Uiedl,T.B., Toshilyuk,I.S.,  
 Carinini,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,  
 Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,  
 McEwan,K.J., Malek,J.A., Gunatane,P.H., Richards,S.,  
 Motley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulik,S.W.,  
 Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,  
 Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,  
 Sanchez,A., Whiting,M., Madan,A., Young,A.C., Scherchenko,Y.,  
 Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,  
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmitt,J., Myers,R.M.,  
 Butlerfield,Y.S., Krzywinski,M.I., Skalska,U., Smalins,D.E.,  
 Schnerich,A., Schein,J.E., Jones,S.J. and Marra,M.A.  
 Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 12477932  
 2 (bases 1 to 1470)  
 Strausberg,R.  
 Direct Submission  
 Submitted (01-JUN-2004) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
 Tissue Procurement: Invitrogen  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.W.A.G.E. Consortium (LINL)  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland;  
 Web site: <http://www.nisc.nih.gov/>  
 Contact: [nisc\\_mgc@nih.gov](mailto:nisc_mgc@nih.gov)  
 Akhter,N., Ayala,K., Beckstrom-Sternberg,S.M., Benjamin,B.,  
 Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,

Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,  
 Hansen,N., Ho,S.-L., Karilins,E., Kwong,P., Latic,P., Legaspi,R.,  
 Maduro,O.L., Masfello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,  
 McDowell,J., Pearson,R., Stancirpop,S., Thomas,P.J., Touchman,J.W.,  
 Turgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,  
 Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>  
 Series: IRAC Plate: 174 Row: A Column: 5  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: Hexamer frequency ORF  
 analysis.

## FEATURES

## source

## Location/Qualifiers

1..1470  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5587705"  
 /issue\_type="Ovary, pooled from 3 adults"  
 /clone\_id="NIH MGC\_125"  
 /lab\_host="DH10B"  
 /note="vector: pCMV-SPORT6"  
 <1..1470  
 /gene="TRGV9"  
 /note="synonyms: TCRGV9, V2"  
 /db\_xref="IMGT/LiGM:TRGV9"  
 /db\_xref="LOCUSID:6983"  
 <1..899  
 /gene="TRGV9"  
 /codon\_start=3  
 /product="TRGV9 protein"  
 /protein\_id="AAH2387.1"  
 /db\_xref="GI:47938918"  
 /db\_xref="IMGT/LiGM:TRGV9"  
 /db\_xref="LOCUSID:6983"  
 /db\_xref="LOCUSID:6983"  
 /translation="CVYGAGHLEDPQISSTTSLKRLAEVSGITISATSYWYRE  
 LRGVIVFLVSIYSDGTVRKESGIPSGFEVDRIPESTSLTIHNEKQDIATYCA  
 LMEVQSPWYVILKLGSGITLVTDKQDADVSPRTFLPSIAITKLOKAGTYCLE  
 KFPEDVIKIHQEKSKNTILQSGQNTMTKMTKPSMLTPPKSLDKERHCIVRHE  
 NKNKVQOEIIFPKITDVTIMDPKDKSKDANDTLQLTNTSAYWYVLLLLKSVV  
 YFAITTCILKRTAFCNGERS"

## CDS

## gene

## ORIGIN

## Alignment Scores:

Pred. No.: 1,49e-29 Length: 1470  
 Score: 307.00 Matches: 58  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-10-031-158-14 (1-58) x BC072387 (1-1470)

Qy 1 MetGlnMetPheProPheSerProLeuPhePheLeuGlnLeuLeuLyGlnSerSer 20  
 Db 391 ATGCAGATGTTCCCGCCAGCCCACTATTTTCTTCCTCAATGCTGAAACCAAGCTCC 450

Qy 21 ArgArgLeuGlnHisThrPheValPheLeuArgAenPheSerLeuMetLeuLeuArgTyr 40  
 Db 451 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTATTAAGATAC 510

Qy 41 lIleGlyLysLysArgArgAlaThrArgPheTrpApeProArgArgGlyThrPro 58  
 Db 511 ATTGGCAAGAAAGAAAGCAACGATTCGTGGATTCAGAGGAGGAAACACCA 564

RESULT 22  
 HMMTCRGA 1530 bp mRNA linear PRI 14-JAN-1995  
 LOCUS Human T-cell receptor gamma chain VCII-CII region mRNA,  
 DEFINITION complete cds.  
 ACCESSION M16768  
 VERSION M16768.1 GI:339399

KEYWORDS T-cell receptor.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 1530)  
AUTHORS Krangel, M.S., Band, H., Hata, S., McLean, J., and Brenner, M.B.  
TITLE Structurally divergent human T cell receptor gamma proteins encoded by distinct C gamma genes  
JOURNAL Science 237 (4810), 64-67 (1987)  
MEDLINE 87263363  
PUBMED 2955517  
COMMENT Original source text: Human (cell line PBL Cl), cDNA to mRNA, clone PBLCl.15.  
Only a few V and J region genes occur within the genome. Diversity in T-cell receptors occur in the V-J recombination events. In some cases, this produces unproductively rearranged reading frames. Also, in some recombination events, some gene regions can be included more than one time creating even more diversity in TCR's.

FEATURES  
source location/qualifiers  
1..1530  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/map="7p15-p14"  
1..1530  
/gene="TCRG"  
/gene="TCRG"  
/product="TCR-gamma mRNA"  
153..1085  
/gene="TCRG"  
/note="T-cell receptor (V-J-C) precursor"  
/codon\_start=1  
/protein\_id="AA61110.1"  
/db\_xref="GI:339400"  
/db\_xref="GDB:G00-120-407"  
translation="MSLHASTAVAGALCVGAGHLBOPISSTTKLSKTRLECV  
VSGTSATSVYWRERPGVIOPLVSIISIDGTAKSGSPSGFEVDRIPESTSLV  
TINAVEKODIATYICALBEGYKLFPSGTTLVTDQOLDVSPKTFILPSTAEK  
LQKAGTYLCLEKPPVDVIRIHWQEKSNILSGQENKTMDTYKFFSWLTPESK  
LDKSHRCIVHNNKNGVDQELIFPIKTYITWDPDNCSKDANDTLILQITVTSAY  
YMYLLLSKSVYFAITTCILRTATACNGEKS"  
153..197  
/gene="TCRG"  
/note="T-cell receptor signal peptide"  
198..1082  
/gene="TCRG"  
/product="T-cell receptor (V-J-CI-CII)"  
517..518  
/gene="TCRG"  
/organism="Homo sapiens"  
564..565  
/gene="TCRG"  
/organism="Homo sapiens"  
/gene="TCRG"  
894..895  
/gene="TCRG"  
/organism="Homo sapiens"  
942..943  
/gene="TCRG"  
/organism="Homo sapiens"  
ORIGIN Chromosome 7p15.  
Alignment Scores:  
Score: 1.55e-29 Length: 1530  
Percent Similarity: 307.00 Matches: 58  
Best Local Similarity: 100.00% Conservative: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0  
US-10-031-158-14 (1-58) x HUMTCRGA (1-1530)

Qy 1 MetGlnMetPheProSerProLeuPhePhePheLeuGlnLeuLeuTyGlnSerSer 20  
Db 577 ATGAGATGTTTCCCAAGCCCACTATTTCTTCCTCAATGCTGAAACAAAGCTCC 636  
Qy 21 ArgArgLeuGlnHisThrPheValPheLeuArgPheSerLeuMetLeuLeuArgTyr 40  
Db 637 AGAAGCTGGAACATACCTTGTCTTGAAGAAATTTTCCCGATGTTATTAAGATAC 696  
Qy 41 ILGLVLYSLYSARGALATNRARGPheTTPAsPProARARGSLYTHRPro 58  
Db 697 ATTGGCAAGAAAGAGACACGATCTGGATCCAGAGGGAGACCA 750

RESULT 23  
LOCUS BC072396  
DEFINITION Homo sapiens T cell receptor gamma variable 9, mRNA (cDNA clone MGC:90486 IMAGE:575352), complete cds.  
ACCESSION BC072396  
VERSION BC072396.1 GI:47939858  
KEYWORDS MGC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 1825)  
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klauener, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altshul, S.F., Zeeberg, B., Bluetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.L., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Ushed, T.B., Tothiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loggellano, N.A., Peters, G.J., Abramson, R.D., Mulvaney, S.J., Bosak, S.A., McEwen, P.J., McKernan, K.J., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kesteman, M., Madan, A., Young, A.C., Shvchenko, Y., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shvchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marz, M.A.  
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
12477932  
2 (bases 1 to 1825)  
Strausberg, R.  
Direct Submission  
Submitted (01-JUN-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
NIH-MGC Project URL: http://mgc.ncl.nih.gov  
Contact: MGC help desk  
Email: gcgaps-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)  
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),  
Gaithersburg, Maryland  
Web site: http://www.nisc.nih.gov/  
Contact: nisc\_mgc@nigr.nih.gov  
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Hachighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kong, P., Latic, P., Legaas, R., Maduro, Q.L., Maestello, C., Maekeri, B., Maestri, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stancirip, S., Thomas, P.J., Touchman, J.W., Tsurgoun, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILMIL ac: <http://image.llnl.gov>  
 Series: IRAX Plate: 174 Row: c Column: 15  
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

## FEATURES

source  
 Location/Qualifiers  
 1. 1825  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="MGC:90486 IMAGE:575352"  
 /tissue\_type="Blood, adult leukocytes"  
 /clone\_id="NIH\_MGC\_118"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"  
 1. 1825  
 /gene="TRGV9"  
 /note="synonyms: TCRGV9, V2"  
 /db\_xref="IMGT/TCR.V9"  
 /db\_xref="IMGT/TCR.V9"  
 378. 1325  
 /gene="TRGV9"  
 /codon\_start=1  
 /product="TRGV9 protein"  
 /protein\_id="AAH72396.1"  
 /db\_xref="GI:47939859"  
 /db\_xref="IMGT/TCR.V9"  
 /db\_xref="LocusID:6983"  
 /translation="MLSLHTSTLAVALGALCVYAGHLEPOQISTSTLTSTANLBCV  
 VSGITISATVYWRERGEVIOPLVSIISYDGTAKESGIPSGREVDRIPEITSTLT  
 TIHWVKODIATYLCALWEVGEGLKIKVFGPGTKLIITDKDLADVSPPTIFLPS  
 IAEFLQKAGTYLCLLEKFPDVKIKHWEKSKNTIIASOGAMTKNDTDMKFSMLT  
 VPEKSIDENRCHVHNNKNGVDOELIIFPKIDVITMDKXNDSKXANTLLQLT  
 NTSAYTYLLILIKSVYFAITTCILRLRTAFCCNGEKS"

## CDS

## gene

## ORIGIN

## Alignment Scores:

Pred. No.:	1.89e-29	Length:	1825
Score:	307.00	Matches:	58
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-10-031-158-14 (1-58) x BC072396 (1-1825)

OY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuLeuGlnSerSer 20  
 DB 817 ATGCAGATGTTTCCCCCAAGCCCACTATTTCTTCTCAATTCCTGAAACAAAGCTCC 876  
 OY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40  
 DB 877 AGAAGCGTGAACATACCTTGTCTCTTGTGAGAAAATTTTCCCTAATGTTATTAAAGATAC 936  
 OY 41 ILeGlyLysLysArgArgGlnThrArgPheTyrPheProArgArgGlyThrPro 58  
 DB 937 ATTGGAGAAAGAAAGACCAACGATTCCTGGATCCAGAGAGGGAACACCA 990

## RESULT 24

AC130306 66558 bp DNA linear HTG 09-AUG-2002  
 LOCUS AC130306 Homo sapiens chromosome 11 clone RP11-126H19 map 11, LOW-PASS  
 DEFINITION SEQUENCE SAMPLING.

ACCESSION AC130306  
 VERSION AC130306.1 GI:22165240  
 KEYWORDS HTG; HTGS PHASE0.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE 1 Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 66558)

AUTHORS  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORSTITLE  
 JOURNAL  
 COMMENT

Birtren,B., Nusbaum,C. and Lander,E.  
 Homo sapiens chromosome 11, clone RP11-126H19  
 Unpublished  
 2 (bases 1 to 66558)  
 Birtren,B., Nusbaum,C., Lander,E., All,A., Allen,N., Anderson,S.,  
 Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B.,  
 Camarata,J., Cheng,J., Chazaro,B., Choepel,Y., Collymore,A.,  
 Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
 Fero,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,  
 Gardaya,S., Gord,S., Graham,L., Grand-Pierre,N., Hages,B.,  
 Horton,L., Hulme,M., Iliev,I., Johnson,R., Jones,C., Kamat,A.,  
 Karas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,  
 Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,  
 McCarthy,M., Melidrim,J., Meneus,L., Milnova,T., Menga,V.,  
 Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Notman,C.H.,  
 O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,  
 Puhunkang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,  
 Roman,J., Roy,A., Schauer,S., Schupack,R., Seaman,S., Severy,P.,  
 Smith,C., Spencer,B., Strange-Thomann,N., Stojanovic,N., Talmage,J.,  
 Teafaye,S., Theodore,J., Topnam,K., Travers,M., Vassiliev,H.,  
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,  
 Zembek,L., Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted (09-AUG-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
 ----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: <http://www-seq.wi.mit.edu>  
 Contact: [sequence.submissions@genome.wi.mit.edu](mailto:sequence.submissions@genome.wi.mit.edu)  
 ----- Project Information  
 Center project name: L27898  
 Center clone name: 126\_H\_19  
 -----  
 \* NOTE: This record contains 80 individual  
 \* sequencing reads that have not been assembled into  
 \* contigs. Runs of N are used to separate the reads  
 \* and the order in which they appear is completely  
 \* arbitrary. Low-pass sequence sampling is useful for  
 \* identifying clones that may be gene-rich and allows  
 \* overlap relationships among clones to be deduced.  
 \* However, it should not be assumed that this clone  
 \* will be sequenced to completion. In the event that  
 \* the record is updated, the accession number will  
 \* be preserved.  
 1 735: contig of 735 bp in length  
 736 835: gap of 100 bp  
 836 1572: contig of 737 bp in length  
 1573 1672: gap of 100 bp  
 1673 2399: contig of 727 bp in length  
 2400 2499: gap of 100 bp  
 2500 3220: contig of 721 bp in length  
 3221 3320: gap of 100 bp  
 3321 4037: contig of 717 bp in length  
 4038 4137: gap of 100 bp  
 4138 4881: contig of 744 bp in length  
 4882 4981: gap of 100 bp  
 4982 5714: contig of 733 bp in length  
 5715 5814: gap of 100 bp  
 5815 6541: contig of 727 bp in length  
 6542 6641: gap of 100 bp  
 6642 7361: contig of 720 bp in length  
 7362 7461: gap of 100 bp  
 7462 8214: contig of 753 bp in length  
 8215 8314: gap of 100 bp  
 8315 9056: contig of 742 bp in length  
 9057 9156: gap of 100 bp  
 9157 9916: contig of 760 bp in length  
 9917 10016: gap of 100 bp  
 10017 10732: contig of 716 bp in length

10733 10832: gap of 100 bp  
10833 11543: contig of 711 bp in length  
11544 11643: gap of 100 bp  
11644 12392: contig of 749 bp in length  
12393 12492: gap of 100 bp  
12493 13224: contig of 732 bp in length  
13225 13324: gap of 100 bp  
13325 14045: contig of 721 bp in length  
14046 14145: gap of 100 bp  
14146 14894: contig of 749 bp in length  
14895 14994: gap of 100 bp  
14995 15726: contig of 732 bp in length  
15727 15826: gap of 100 bp  
15827 15662: contig of 736 bp in length  
15663 16662: gap of 100 bp  
16663 17400: contig of 738 bp in length  
17401 17500: gap of 100 bp  
17501 18229: contig of 729 bp in length  
18230 18329: gap of 100 bp  
18330 19041: contig of 712 bp in length  
19042 19141: gap of 100 bp  
19142 19860: contig of 719 bp in length  
19861 19960: gap of 100 bp  
19961 20680: contig of 720 bp in length  
20681 20780: gap of 100 bp  
20781 21508: contig of 728 bp in length  
21509 21608: gap of 100 bp  
21609 22316: contig of 708 bp in length  
22317 22416: gap of 100 bp  
22417 23158: contig of 742 bp in length  
23159 23258: gap of 100 bp  
23259 24028: contig of 770 bp in length  
24029 24128: gap of 100 bp  
24129 24867: contig of 739 bp in length  
24868 24967: gap of 100 bp  
24968 25695: contig of 728 bp in length  
25696 25795: gap of 100 bp  
25796 26518: contig of 723 bp in length  
26519 26618: gap of 100 bp  
26619 27359: contig of 741 bp in length  
27360 27459: gap of 100 bp  
27460 28200: contig of 741 bp in length  
28201 28300: gap of 100 bp  
28301 29009: contig of 709 bp in length  
29010 29109: gap of 100 bp  
29110 29843: contig of 734 bp in length  
29844 29943: gap of 100 bp  
29944 30683: contig of 740 bp in length  
30684 30783: gap of 100 bp  
30784 31517: contig of 734 bp in length  
31518 31617: gap of 100 bp  
31618 32351: contig of 734 bp in length  
32352 32451: gap of 100 bp  
32452 33177: contig of 726 bp in length  
33178 33277: gap of 100 bp  
33278 33982: contig of 715 bp in length  
33983 34092: gap of 100 bp  
34093 34819: contig of 727 bp in length  
34820 34919: gap of 100 bp  
34920 35636: contig of 717 bp in length  
35637 35736: gap of 100 bp  
35737 36449: contig of 713 bp in length  
36450 36549: gap of 100 bp  
36550 37290: contig of 741 bp in length  
37291 37390: gap of 100 bp  
37391 38129: contig of 739 bp in length  
38130 38229: gap of 100 bp  
38230 38862: contig of 733 bp in length  
38863 39062: gap of 100 bp  
39063 39798: contig of 736 bp in length  
39799 39898: gap of 100 bp  
39899 40608: contig of 710 bp in length  
40609 40708: gap of 100 bp

40709 41454: contig of 746 bp in length  
41455 41554: gap of 100 bp  
41555 42312: contig of 758 bp in length  
42313 42412: gap of 100 bp  
42413 43160: contig of 748 bp in length  
43161 43260: gap of 100 bp  
43261 43996: contig of 736 bp in length  
43997 44096: gap of 100 bp  
44097 44825: contig of 730 bp in length  
44826 44925: gap of 100 bp  
44927 45661: contig of 735 bp in length  
45662 45761: gap of 100 bp  
45762 46504: contig of 743 bp in length  
46505 46604: gap of 100 bp  
46605 47333: contig of 729 bp in length  
47334 47433: gap of 100 bp  
47434 48167: contig of 734 bp in length  
48168 48267: gap of 100 bp  
48268 48994: contig of 727 bp in length  
48995 49094: gap of 100 bp  
49095 49842: contig of 748 bp in length  
49843 49942: gap of 100 bp  
49943 50670: contig of 728 bp in length  
50671 50770: gap of 100 bp  
50771 51509: contig of 739 bp in length  
51510 51610: gap of 100 bp  
51610 52357: contig of 747 bp in length  
52357 52456: gap of 100 bp  
52457 53183: contig of 727 bp in length  
53184 53283: gap of 100 bp  
53284 54003: contig of 720 bp in length  
54004 54103: gap of 100 bp  
54104 54830: contig of 727 bp in length  
54831 54930: gap of 100 bp  
54931 55685: contig of 755 bp in length  
55686 55785: gap of 100 bp  
55785 56514: contig of 729 bp in length  
56514 56614: gap of 100 bp  
56615 57347: contig of 733 bp in length  
57348 57447: gap of 100 bp  
57448 58165: contig of 718 bp in length  
58166 58265: gap of 100 bp  
58265 59004: contig of 739 bp in length

## Alignment Scores:

Pred. No.: 1.05e-27 Length: 66558  
Score: 307.00 Matches: 58  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 2 Gaps: 0

US-10-031-158-14 (1-58) x AC130306 (1-66558)

Qy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20

Db 22925 ATGCAGATGTTTCCCGCAGCCACTATTTCTTCTTCATTCGCAACCAAGCTCC 22984

Qy 21 ArgArgLeuGlnIuriIsthrPheValPheLeuArgAnPheSerLeuMetLeuArgTyr 40

Db 22985 AGAAGGCTGGAACATCTTGTCTTCTTGGAATTTTCCGATGTTATTAAGATAC 23044

Qy 41 TLeGlyLysLysAlaGArgAlaThrArgPheTrrPheProAlaGArgGlyThrPro 58

Db 23045 ATTGGCAAGAAAGAAAGCAACAGATTCGGGATCCAGAGAGGAACACCA 23098

## RESULT 25

AF159056 140691 bp DNA linear PRI 17-AUG-2001  
LOCUS Homo sapiens T-cell gamma receptor locus, complete sequence.  
DEFINITION AF159056  
ACCESSION AF159056  
VERSION AF159056.1 GI:5566238  
KEYWORDS



```

/note="V6P"
/pseudo
complement(78838,.79000)
/rpt_family="MIR"
/rpt_type=dispersed
complement(80629,.81096)
/rpt_family="L1"
/rpt_type=dispersed
8190,.82202
/note="Y4P"
/pseudo
complement(82530,.83651)
/rpt_family="LTR1"
/rpt_type=dispersed
complement(82876,.83159)
/rpt_family="Alu"
/rpt_type=dispersed
85826,.86088
/rpt_family="MER2"
/rpt_type=dispersed
86291,.86368
/rpt_family="MIR"
/rpt_type=dispersed
86542,.86631
/rpt_type=dispersed
/rpt_family="MERS"
/rpt_type=dispersed
complement(90148,.90424)
/rpt_family="Alu"
/rpt_type=dispersed
97655,.97996
/note="JPI"
complement(99475,.99769)
/rpt_family="Alu"
/rpt_type=dispersed
100267,.100327
/note="JP"
complement(103808,.104066)
/rpt_family="Alu"
/rpt_type=dispersed
104364,.104423
/note="J1"
107293,.107592
/rpt_family="Alu"
/rpt_type=dispersed
108009,.108085
/rpt_family="MR"
/rpt_type=dispersed
108237,.108566
/note="C1 exon 1"

Alignment Scores:
Pred. No.:          2_42e-27      Length:           140691
Score:              307.00        Matches:           58
Percent Similarity: 100.00%       Conservative:     0
Best Local Similarity: 100.00%    Mismatches:      0
Query Match:        100.00%       Indels:          0
DB:                 9             Gaps:            0

US-10-031-158-14 (1-58) x AF159056 (1-140691)
QY      1 MetGlnMetPheProSerProLeuPhePhePheLeuGlnLeuLeuYgInserSer 20
Db      108249 ATGCAGATGTTTCCCCCAAGCCACATATTTCCTCTCAATTGGCGAAACAAGCTCC 108306
QY      21 ARGATGleuGlnHisThrPheValPheLeuARGAsnPheserLeuMetLeuLeuArgTyr 40
Db      108309 AGAAGGCTGGAACATAACCCTTGCTCTTGAGAATAATTTTTCCCTGATGTATTAAAGTAC 108366
QY      41 IlegIyLyVsArGrArGlaTrArGrPheTrpAspProArGrArglyThrPro 58
Db      108369 ATTGGCAAGAAAAGAAAGACACACATCTTGGGATCTCCAGAGGGGAACACCA 108422

```

LOCUS	AC006033	171816 bp	DNA	linear	PRI 31-JAN-2004
DEFINITION	Homo sapiens BAC clone RP11-121A8 from 7, complete sequence.				
ACCESSION	AC006033				
VERSION	AC006033.2 GI:4309948				
KEYWORDS	HTG.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	<p>1 (bases 1 to 171816)  Hillier L.M., Fulton R.S., Fulton L.A., Graves T.A., Pepin K.H., Wagner-McPherson C., Layman D., Maas J., Jaeger S., Walker R., Wyile K., Seaton M., Becker M.C., O'Laughlin M.D., Schaller M.E., Fowell G.A., Delahunty K.D., Minet T.L., Nash W.E., Cordes M., Du H., Sun H., Edwards J., Bradshaw-Cordum H., Ali J., Andrews S., Jask A., Vanbrunt A., Nguyen C., Du F., Lamar B., Courtney L., Kalicki J., Ozersky P., Bielicki L., Scott K., Holmes A., Harkins R., Harris A., Strong C.M., Hou S., Tomlinson C., Dauphin-Kohlberg S., Kozlowicz-Reilly A., Leonard S., Kohling T., Rock S.M., Tin-Wollam A.M., Abbott A., Minx P., Maun J., Stromatac C., Latreille P., Miller N., Johnson D., Murray J., Moesner J.P., Wendl M.C., Yang S.P., Schultz B.R., Wallis J.W., Spieth J., Bieri T.A., Nelson J.O., Berkowicz N., Kohlmann P.E., Cook L.L., Hickenbotham M.T., Eldred J., Williams D., Bedell D.A., Matsis E.R., Gillett S.W., Chissoe S.L., Marra M.A., Raymond C., Haugen E., Gillett M., Zhou Y., James R., Phelps K., Iadonoto S., Puhb K., Simms E., Levy R., Clendenning J., Kaul R., Kent W.J., Furey T.S., Baertsch R.A., Brent M.R., Kelbler E., Flick P., Bork P., Suyama M., Bailey J.A., Portnoy M.E., Torrente D., Chinwalla A.T., Ghosh W.R., Eddy S.R., McPherson J.D., Olson M.V., Eichler E.E., Green E.D., Waterston R.H. and Wilson R.K.</p> <p>The DNA sequence of human chromosome 7</p> <p>Nature 424 (6945), 157-164 (2003)</p>				
TITLE	22737999				
JOURNAL	12853948				
MEDLINE	2 (bases 1 to 171816)				
PUBMED	Arnett C., Le T.P. and Wohlmann P.				
REFERENCE	The sequence of Homo sapiens BAC clone RP11-121A8				
AUTHORS	Unpublished (2001)				
TITLE	3 (bases 1 to 171816)				
JOURNAL	Waterston R.H.				
AUTHORS	Direct Submission				
TITLE	Submitted (22-NOV-1998) Genome Sequencing Center, Washington				
JOURNAL	University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA				
REFERENCE	4 (bases 1 to 171816)				
AUTHORS	Waterston R.				
TITLE	Direct Submission				
JOURNAL	Submitted (27-FEB-1999) Department of Genetics, Washington				
AUTHORS	University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA				
REFERENCE	5 (bases 1 to 171816)				
AUTHORS	Waterston R.				
TITLE	Direct Submission				
JOURNAL	Submitted (21-DEC-1999) Department of Genetics, Washington				
AUTHORS	University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA				
REFERENCE	6 (bases 1 to 171816)				
AUTHORS	Waterston R.				
TITLE	Direct Submission				
JOURNAL	Submitted (29-APR-2003) Department of Genetics, Washington				
AUTHORS	University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA				
REFERENCE	7 (bases 1 to 171816)				
AUTHORS	Wilson R.				
TITLE	Direct Submission				
JOURNAL	Submitted (31-JUN-2004) Department of Genetics, Washington				
AUTHORS	University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA				
COMMENT	On Mar 1, 1999 this sequence version replaced gi:3907471.				
	----- Genome Center				
	Center: Washington University Genome Sequencing Center				
	Catalog code: WUGSC				
	Web site: http://genome.wustl.edu				
	Contact: bapens@wustl.wustl.edu				
	----- Summary Statistics				

Center project name: H\_NH0121A08

-----  
 NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clones sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

#### MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GRB/CHR7>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu>

#### SOURCE INFORMATION:

The RPc1-11 human BAC library was made from the blood of one male donor, as described by Ooegawa, K., Moon, P.Y., Zhao, B., Frengen, E., Tatenno, M., Catanese, J.V., and de Jong, P.V. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pBAC3.6

#### NEIGHBORING SEQUENCE INFORMATION:

Actual start of this clone is at base position 1 of RP11-121A8  
 actual end is at 171816 of RP11-121A8.

#### FEATURES

##### SOURCE

```
1..171816
  /organism="Homo sapiens"
  /mol_type="Genomic DNA"
  /db_xref="taxon:9606"
  /chromosome="7"
  /map="7"
  /clone="RP11-121A8"
  /clone_11b="RPc1-11"
  446..928
    repeat_region
      /rpt_family="L2"
    1675..1829
      repeat_region
        /rpt_family="MIR"
    2729..2887
      repeat_region
        /rpt_family="L2"
    3315..3420
      repeat_region
        /rpt_family="MIR"
    3941..3976
      repeat_region
        /rpt_family="AT_rich"
    4114..4171
      repeat_region
        /rpt_family="MIR"
    4175..4305
      repeat_region
        /rpt_family="MER1_type"
    4623..4781
      repeat_region
        /rpt_family="MER2_type"
    4889..5191
      repeat_region
        /rpt_family="L2"
    5209..5499
      repeat_region
        /rpt_family="MALR"
    5512..5623
      repeat_region
        /rpt_family="(TA)n"
    7420..7746
      repeat_region
        /rpt_family="MER2_type"
    8185..8258
      repeat_region
```

```
repeat_region
  /rpt_family="L2"
  8462..9500
    repeat_region
      /rpt_family="L2"
      10020..10040
        repeat_region
          /rpt_family="(TG)n"
          10941..11048
            repeat_region
              /rpt_family="MER2_type"
              11050..11275
                repeat_region
                  /rpt_family="MER2_type"
                  11304..11380
                    repeat_region
                      /rpt_family="(TATG)n"
                      11436..11538
                        repeat_region
                          /rpt_family="L1"
                          11586..12300
                            repeat_region
                              /rpt_family="L1"
                              12313..13051
                                repeat_region
                                  /rpt_family="L1"
                                  14179..14497
                                    repeat_region
                                      /rpt_family="Alu"
                                      14583..14644
                                        repeat_region
                                          /rpt_family="MIR"
                                          15408..15437
                                            repeat_region
                                              /rpt_family="AT_rich"
                                              15974..16255
                                                repeat_region
                                                  /rpt_family="Alu"
                                                  17692..17733
                                                    repeat_region
                                                      /rpt_family="AT_rich"
                                                      18033..18147
                                                        repeat_region
                                                          /rpt_family="(TTTC)n"
                                                          18148..18401
                                                            repeat_region
                                                              /rpt_family="Alu"
                                                              19188..19418
                                                                repeat_region
                                                                  /rpt_family="MER1_type"
                                                                  19779..19908
                                                                    repeat_region
                                                                      /rpt_family="MALR"
                                                                      20266..20321
                                                                        repeat_region
                                                                          /rpt_family="L2"
                                                                          22450..22481
                                                                            repeat_region
                                                                              /rpt_family="AT_rich"
                                                                              22579..22887
                                                                                repeat_region
                                                                                  /rpt_family="Alu"
                                                                                  23438..23475
                                                                                    repeat_region
                                                                                      /rpt_family="AT_rich"
                                                                                      23476..23944
                                                                                        repeat_region
                                                                                          /rpt_family="L1"
                                                                                          26364..26491
                                                                                            repeat_region
                                                                                              /rpt_family="L1"
                                                                                              26619..26798
                                                                                                repeat_region
                                                                                                  /rpt_family="L1"
                                                                                                  27125..27240
                                                                                                    repeat_region
                                                                                                      /rpt_family="MALR"
                                                                                                      27412..27433
                                                                                                        repeat_region
                                                                                                          /rpt_family="AT_rich"
                                                                                                          27672..27685
                                                                                                            repeat_region
                                                                                                              /rpt_family="Alu"
```

#### Alignment Scores:

```
Pred. No.: 3.02e-27 Length: 171816
Score: 307.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
```

US-10-031-158-14 (1-58) x AC006033 (1-171816)

```
Qy 1 MetGlnMetPheProPheSerProLeuPhePheLeuGlnLeuLeuLeuGlnSerSer 20
Db 86565 ATGCAGATGTTCCCAAGCCCACTATTTCTTCCTTCATTCGTGAACAAAGCTCC 86506
Qy 21 ATGATGLeuGlnLHtThrPheValaPheLeuAArgaPheSerLeuMetLeuLeuArgTyr 40
Db 86505 AGAAGCTGGAACATACCTTGTCTTGTGAAGAAATTTTCCCTGATGTTAAGATAC 86446
```

Qy 41 |||G|L|Y|S|Y|S|A|G|A|T|A|T|H|A|T|H|P|H|E|T|P|A|P|P|P|O|A|G|A|G|G|Y|T|H|P|P|O| 58  
 Db 86445 ATTGGCAAGAAAGAAAGACACACGATTCTGGAGATCCACGAGGAGGAAACACCA 86392

RESULT 27  
 HSTCELG1 400 bp DNA linear PRI 02-DEC-2003  
 DEFINITION Human gene fragment for T-cell receptor gamma 1 chain constant region (loc. on chromosome 7p15).  
 ACCESSION X02766  
 VERSION X02766.1 GI:36779  
 KEYWORDS T-cell receptor.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 Murre, C., Malmann, R.A., Morton, C.C., Bongiovanni, K.F., Malmann, T.A., Shows, T.B. and Seidman, J.G.  
 Human gamma-chain genes are rearranged in leukaemic T cells and map to the short arm of chromosome 7  
 Nature 316 (6028), 549-552 (1985)

JOURNAL MEDLINE 85296267  
 PUBMED 3875797

FEATURES  
 source location/Qualifiers  
 1..400 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 <1..27 /note="put. intron"  
 <28..357 /note="unamed protein product; (g1) C region fragment (357 is 1st base in codon)"  
 /codon\_start=3 /protein\_id="CAE82035.1"  
 /db\_xref="GI:38635935"  
 /translation="KQIDADVSPPRTFLPSIATKIQKACTYICLAEKFPDYIKTH WEERKSNITLGSQSGNTWKNTDVTWKFSWLTVPKSLDKERICTVRHNNKNVDQEI IFPIKTK"  
 358..>400 /note="put. intron"

intron  
 intron

CDS  
 <28..357 /note="unamed protein product; (g1) C region fragment (357 is 1st base in codon)"  
 /codon\_start=3 /protein\_id="CAE82035.1"  
 /db\_xref="GI:38635935"  
 /translation="KQIDADVSPPRTFLPSIATKIQKACTYICLAEKFPDYIKTH WEERKSNITLGSQSGNTWKNTDVTWKFSWLTVPKSLDKERICTVRHNNKNVDQEI IFPIKTK"  
 358..>400 /note="put. intron"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 8.44e-30 Length: 400  
 Score: 304.00 Matches: 57  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 98.28% Mismatches: 0  
 Query Match: 99.02% Indels: 0  
 Gaps: 0

US-10-031-158-14 (1-58) x HSTCELG1 (1-400)

Qy 1 MetGlnMetPheProPheProSerProLeuPhePhePheLeuGlnLeuLeuGlnSerSer 20  
 Db 40 ATGAGAGATGTTCCCGCCAGCCCACTATTTCTTCTTCATATGCTGAAACAAACCTCC 99

Qy 21 ArgArgLeuGlnLuhIstHrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40  
 Db 100 AGAAGGCTGGAACATACCTTGTCTCTTGGAATTTTTCCTCGATATTATTAAGATAC 159

Qy 41 |||G|L|Y|S|Y|S|A|G|A|T|A|T|H|A|T|H|P|H|E|T|P|A|P|P|P|O|A|G|A|G|G|Y|T|H|P|P|O| 58  
 Db 160 ATTGGCAAGAAAGAAAGACACACGATTCTGGAGATCCACGAGGAGGAAACACCA 213

RESULT 28  
 CQ493135/c CQ493135 486 bp DNA linear PAT 30-JAN-2004  
 LOCUS Sequence 25002 from Patent WO0160860.  
 DEFINITION CQ493135  
 ACCESSION CQ493135  
 VERSION CQ493135.1 GI:41458754

KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 AUTHORS Schlegel, R., Endege, W.O. and Monahan, J.E.  
 TITLE Genes differentially expressed in human prostate cancer and their use  
 JOURNAL Patent: WO 0160860-A 25002 23-AUG-2001; Millennium Predictive Medicine, Inc. (US)  
 FEATURES  
 source location/Qualifiers  
 1..486 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.05e-29 Length: 486  
 Score: 304.00 Matches: 57  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 98.28% Mismatches: 0  
 Query Match: 99.02% Indels: 0  
 Gaps: 0

US-10-031-158-14 (1-58) x CQ493135 (1-486)

Qy 1 MetGlnMetPheProPheProSerProLeuPhePhePheLeuGlnLeuLeuGlnSerSer 20  
 Db 411 ATGAGAGATGTTCCCGCCAGCCCACTATTTCTTCTTCATATGCTGAAACAAACCTCC 352

Qy 21 ArgArgLeuGlnLuhIstHrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40  
 Db 351 AGAAGGCTGGAACATACCTTGTCTCTTGGAATTTTTCCTCGATATTATTAAGATAC 292

Qy 41 |||G|L|Y|S|Y|S|A|G|A|T|A|T|H|A|T|H|P|H|E|T|P|A|P|P|P|O|A|G|A|G|G|Y|T|H|P|P|O| 58  
 Db 291 ATTGGCAAGAAAGAAAGACACACGATTCTGGAGATCCACGAGGAGGAAACACCA 238

RESULT 29  
 CQ144550 CQ144550 316 bp DNA linear PAT 21-JAN-2004  
 LOCUS Sequence 14572 from Patent WO0157276.  
 ACCESSION CQ144550  
 VERSION CQ144550.1 GI:41101922  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.  
 TITLE Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human bone marrow  
 JOURNAL Patent: WO 0157276-A 14572 09-AUG-2001; Aeomica, Inc. (US)  
 FEATURES  
 source location/Qualifiers  
 1..316 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 /note="MAP TO AF159056.1-EXPRESSED IN BONE MARROW, SIGNAL = 2.2-SWISSPROT HIT: P03986, EVALU6 7.00e-58-NT HIT: M14996.1, EVALU6 0.00e+00-EST\_HUMAN HIT: BF679123.1, EVALU6 0.00e+00"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.17e-29 Length: 316  
 Score: 302.00 Matches: 57  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0

Query Match: 98.37% Indels: 0  
 DB: 6 Gaps: 0

US-10-031-158-14 (1-58) x CQ204403 (1-316)

QY 2 GlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYegInSerSerArg 21  
 Db 2 CAGATGTTCCCAAGCCCACTATTCTTCTTCCTCAATTCCTGAACCAAGCTCCAGA 61

QY 22 ArgLeuGlnIsthrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyrIle 41  
 Db 62 AGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAGATACATT 121

QY 42 GlyIysIysArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58  
 Db 122 GGCAGAGAAAGAGAGCAACGATTCTGGATCCCAAGAGGAGAACCA 172

RESULT 30  
 LOCUS CQ204403 316 bp DNA linear PAT 21-JAN-2004  
 DEFINITION Sequence 6680 from Patent WO0157271.  
 ACCESSION CQ204403  
 VERSION CQ204403.1 GI:41190250  
 KEYWORDS  
 ORGANISM Homo sapiens (human)  
 SOURCE Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE  
 1 Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.  
 Human genome-derived single exon nucleic acid probes useful for  
 analysis of gene expression in human breast and bc 474 cells  
 Patent: WO 0157271-A 6680 09-AUG-2001;  
 Aecmics, Inc. (US)  
 Location/Qualifiers  
 source  
 1..316  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 /note="MAP TO AF159056.1-EXPRESSED IN BT474, SIGNAL =  
 1.1-SWISSPROT HIT: P03986, EVALUE 7.00e-58-NT HIT:  
 M14996.1, EVALUE 0.00e+00-EST\_HUMAN HIT: Bf679123.1,  
 EVALUE 0.00e+00"

ORIGIN

Alignment Scores:  
 Pred. No.: 1,17e-29 Length: 316  
 Score: 302.00 Matches: 57  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 98.37% Indels: 0  
 DB: 6 Gaps: 0

US-10-031-158-14 (1-58) x CQ204403 (1-316)

QY 2 GlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYegInSerSerArg 21  
 Db 2 CAGATGTTCCCAAGCCCACTATTCTTCTTCCTCAATTCCTGAACCAAGCTCCAGA 61

QY 22 ArgLeuGlnIsthrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyrIle 41  
 Db 62 AGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAGATACATT 121

QY 42 GlyIysIysArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58  
 Db 122 GGCAGAGAAAGAGAGCAACGATTCTGGATCCCAAGAGGAGAACCA 172

RESULT 31  
 LOCUS CQ302987 316 bp DNA linear PAT 23-JAN-2004  
 DEFINITION Sequence 14092 from Patent WO0186003.  
 ACCESSION CQ302987  
 VERSION CQ302987.1 GI:41263564

KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE  
 1 Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.  
 Human genome-derived single exon nucleic acid probes useful for  
 analysis of gene expression in human lung  
 Patent: WO 0186003-A 14092 15-NOV-2001;  
 Aecmics, Inc. (US)  
 Location/Qualifiers  
 source  
 1..316  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 /note="MAP TO AF159056.1-EXPRESSED IN LUNG, SIGNAL =  
 1.1-SWISSPROT HIT: P03986, EVALUE 7.00e-58-NT HIT:  
 M14996.1, EVALUE 0.00e+00-EST\_HUMAN HIT: Bf679123.1,  
 EVALUE 0.00e+00"

ORIGIN

Alignment Scores:  
 Pred. No.: 1,17e-29 Length: 316  
 Score: 302.00 Matches: 57  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 98.37% Indels: 0  
 DB: 6 Gaps: 0

US-10-031-158-14 (1-58) x CQ302987 (1-316)

QY 2 GlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYegInSerSerArg 21  
 Db 2 CAGATGTTCCCAAGCCCACTATTCTTCTTCCTCAATTCCTGAACCAAGCTCCAGA 61

QY 22 ArgLeuGlnIsthrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyrIle 41  
 Db 62 AGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAGATACATT 121

QY 42 GlyIysIysArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58  
 Db 122 GGCAGAGAAAGAGAGCAACGATTCTGGATCCCAAGAGGAGAACCA 172

RESULT 32  
 LOCUS CQ511627/c 569 bp DNA linear PAT 30-JAN-2004  
 DEFINITION Sequence 43494 from Patent WO0160860.  
 ACCESSION CQ511627  
 VERSION CQ511627.1 GI:41477891  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE  
 1 Schlegel, R., Endege, W.O. and Monahan, J.E.  
 Genes differentially expressed in human prostate cancer and their  
 use  
 Patent: WO 0160860-A 43494 23-AUG-2001;  
 Millennium Predictive Medicine, Inc. (US)  
 Location/Qualifiers  
 source  
 1..569  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

ORIGIN

Alignment Scores:  
 Pred. No.: 7,4e-29 Length: 569  
 Score: 298.00 Matches: 57  
 Percent Similarity: 98.28% Conservative: 0  
 Best Local Similarity: 98.28% Mismatches: 1

Query Match: 97.07% Indels: 0  
 DB: 6 Gaps: 0

US-10-031-158-14 (1-58) x CQ511627 (1-569)

Qy 1 MetGlnMetPheProPheProSerProLeuPhePhePheLeuGlnLeuLeuGlnSerSer 20  
 |||||  
 Db 460 ATGCAGATGTTTCCCGCCAGCCCACTATTTTCTTCCATTCCTGAAACAAAGCTCC 401

Qy 21 ArgArgLeuGlnHsthrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40  
 |||||  
 Db 400 AGAAGCTGGAAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTATTAAGATGC 341

Qy 41 IleGlyLysLysArgArgAlaThrArgPheThrPheProAlaGArgGlyThrPro 58  
 |||||  
 Db 340 ATTGGCAGAAAAAGAGACACAGATTCTGGATCCAGAGGGGAAACACCA 287

RESULT 33  
 HUMTCGCG

LOCUS HUMTCGCG 330 bp DNA linear PRI 13-JAN-1995  
 DEFINITION Human T-cell receptor gamma-chain (TCRG2) gene C2-region  
 allele gamma-2b, exon 1.  
 ACCESSION M15002  
 VERSION M15002.1 GI:339083  
 KEYWORDS C-region; T cell receptor gamma-chain; T-cell receptor; germline.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 330)  
 LeFranc,M.P., Forster,A. and Rabbits,T.H.  
 Genetic polymorphism and exon changes of the constant regions of  
 the human T-cell rearranging gene gamma  
 Proc. Natl. Acad. Sci. U.S.A. 83 (24), 9596-9600 (1986)  
 JOURNAL MEDLINE  
 PUBMED 87092294  
 2879283

COMMENT Original source text: Human B-cell polyclonocytic leukemia line  
 D-PL DNA, clone lambda-D7.  
 Entry revisions for [1] kindly provided by M.-P. LeFranc,  
 22-JUN-1998.

FEATURES  
 source location/Qualifiers  
 1..330  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /map="7p15-p14"  
 1..330  
 /gene="TCRCG2"  
 1..330  
 /gene="TCRCG2"  
 /note="G00-120-409"  
 /number=1

ORIGIN 165 bp upstream of BamHI site.

Alignment Scores:  
 Pred. No.: 1.32e-28 Length: 330  
 Score: 294.00 Matches: 55  
 Percent Similarity: 98.28% Conservative: 2  
 Best Local Similarity: 94.83% Mismatches: 1  
 Query Match: 95.77% Indels: 0  
 DB: 9 Gaps: 0

US-10-031-158-14 (1-58) x HUMTCGCG (1-330)

Qy 1 MetGlnMetPheProPheProSerProLeuPhePhePheLeuGlnLeuLeuGlnSerSer 20  
 |||||  
 Db 13 ATGCAGATGTTTCCCGCCAGCCCACTATTTTCTTCCATTCCTGAAACAAAGCTCC 72

Qy 21 ArgArgLeuGlnHsthrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40  
 |||||  
 Db 73 AGAAGCTGGAAACATACCTTGTCTTCTTGAGAAATTTTCCAGATATTAAGATGC 132

Qy 41 IleGlyLysLysArgArgAlaThrArgPheThrPheProAlaGArgGlyThrPro 58

Db 133 ATTGGCAGAAAAAGAGACACACATTTCTGGATCCAGAGGGGAAACACCA 186  
 |||||

RESULT 34  
 HSTRGCS4  
 LOCUS HSTRGCS4 635 bp mRNA linear PRI 20-MAR-1991  
 DEFINITION Human truncated mRNA from TRG gamma gene V(g)8-J(g)P2-C(g)2 with  
 nonproductive V-J arrangement.  
 ACCESSION X06776  
 VERSION X06776.1 GI:37343  
 KEYWORDS constant region; joining region; T-cell receptor; T-cell receptor  
 gamma; variable region.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 635)  
 Tighe,L., Forster,A., Clark,D.M., Boylston,A.W., Lavenit,I. and  
 Rabbits,T.H.  
 Unusual forms of T cell gamma mRNA in a human T cell leukemia cell  
 line: implications for gamma gene expression  
 Eur. J. Immunol. 17 (12), 1729-1736 (1987)  
 JOURNAL MEDLINE  
 PUBMED 88083067  
 2961573

COMMENT pM17c54 is derived from TRG gamma locus allele 2;  
 see X06775 for further allele 2 nonproductive cDNA; see X06774 for  
 MOL717 TRG gamma allele 1 derived productive cDNA.  
 location/Qualifiers  
 1..635  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="pM17c54"  
 /cell\_line="MOLT17"  
 /clone\_1lb="lambda gt10"  
 2..217  
 /note="unnamed protein product; Protein sequence is in  
 conflict with the conceptual translation; V(g)8 region (AA  
 1-72)"  
 /codon\_start=1  
 /protein\_id="CAA29944.1"  
 /db\_xref="GI:1335357"  
 /translation="LIPENAVVTHWYTHQEGKAPQRLYYDYSNRYVLESISREKY  
 HTYASGKSLKFLIENLIERDSEVYCA"  
 218..227  
 /note="out of frame joint V(g)8-J(g)P2"  
 219..635  
 /note="unnamed protein product; Protein sequence is in  
 conflict with the conceptual translation; pot. aberrant  
 translation product"  
 /codon\_start=1  
 /protein\_id="CAA29945.1"  
 /db\_xref="GI:1335358"  
 /translation="PMRWIKTPFAKGRILVTSPPKQDADVSPKRTFLPSIATKLO  
 KAGTYICLEKRPDDIKIKHWKKSITLIGSEGNMKNDYMKRSMWLTPRESID  
 KERRCIYRHNENKNGIDQETIFPPITKMLIKKKK"  
 228..276  
 /note="J(g)P2 region (truncated at 5'end)"  
 277..605  
 /note="C(g)2 (exon 1)"  
 558..563  
 /note="poc. polyA signal"  
 596..601  
 /note="poc. polyA signal"  
 606..635  
 /note="translated intron + polyA sequence"

ORIGIN misc\_feature  
 misc\_feature  
 misc\_feature

Alignment Scores:  
 Pred. No.: 2.73e-28 Length: 635  
 Score: 294.00 Matches: 55  
 Percent Similarity: 98.28% Conservative: 2  
 Best Local Similarity: 94.83% Mismatches: 1



translation of the corresponding C\_region. Presently translation qualifiers on C\_region features are illegal."

/codon\_start=1  
/protein\_id="AA6112.1"  
/db\_xref="GI:540459"  
/translation="DKQADAVSPKPTFLPSIAETKQKAGTYLCLEKFPDIIKI  
HMOEKSNITLIGSOEGMTKNDYMFSLTVEESLDKEHRCIVHNNKNGIDOE  
IIFPIKDTVDPKRNVSKDANDVTMDPKDMSKDANDTLQLTNTSAYVWYL  
LLKSVYFAIITCCLGRTRFCNGEKS"

C\_region  
/gene="TCRCG2"  
/product="T cell receptor gamma chain"  
/note="G00-120-409"  
989..994  
polyA\_signal  
/gene="TCRCG2"  
/note="G00-120-409"  
1013  
polyA\_signal  
/gene="TCRCG2"  
/note="G00-120-409"

## ORIGIN

Alignment Scores:  
Pred. No.: 4.6e-28 Length: 1013  
Score: 294.00 Matches: 55  
Percent Similarity: 98.28% Conservative: 2  
Best Local Similarity: 94.83% Mismatches: 1  
Query Match: 95.77% Indels: 0  
DB: 9 Gaps: 0

US-10-031-158-14 (1-58) x HUMTCRGA (1-1013)

Qy 1 MetGlnMetPheProPseProLeuPhePheLeuGlnLeuLeuLyGlnSerSer 20  
Db 14 ATGCGATGTTTCCCGCAAGCCCACTATTTCTTCCTTCATTCGTAACAAACTCC 73  
Qy 21 ArgArgLeuGlnHisIsthPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40  
Db 74 AGAAGCGCTGGAACATACCTTGTCTCTTGAGAAATTTTCCAGATATTATTAAGATAC 133  
Qy 41 TleGlyLysLysArgArgAlaThrArgPheTyrAspProArgArgGlyThrPro 58  
Db 134 ATTGGCAAGAAAGAGAGCAACGATTCGGATCCCAAGAGGGAACACCA 187

## RESULT 37

HSTRGC64 1160 bp mRNA linear PRI 12-SEP-1993  
LOCUS Human mRNA for T-cell rearranging gamma gene (TRG) V(g)8-J(g)2-  
C(g)2.  
X06774

ACCESSION X06774.1 GI:37346  
VERSION Constant region; joining region; T-cell receptor; T-cell receptor  
KEYWORDS gamma; variable region.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1160)  
AUTHORS Tishbe, L., Forster, A., Clark, D.M., Boylston, A.W., Lavenir, I. and  
Rabbitts, T.H.  
TITLE Unusual forms of T cell gamma mRNA in a human T cell leukemia cell  
line: implications for gamma gene expression

JOURNAL Eur. J. Immunol. 17 (12), 1723-1736 (1987)  
MEDLINE 88083067  
PubMed 2961573

COMMENT pm17c64 is derived from allele 1 of the TRG gamma locus; see x06765  
and x06776 for MOL17 TRG gamma cDNAs representing unproductive  
rearrangements derived from second allele of the TRG gamma locus.

FEATURES  
source  
1..1160  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="pm17c64"

## COS

/cell\_line="MOLT17"  
/clone\_id="lambda gcl0"  
116..1138  
/note="unnamed protein product; TRG gamma chain (AA  
1-340)"  
/codon\_start=1  
/protein\_id="CAA29941.1"  
/db\_xref="GI:37347"  
/translation="MLALALILALPLPPASQKSNLEGRRTKSVTRPGSSAVITCDLP  
VENAVYTHWTLHDEGRAPQRLIYDSNSRNVLESQISREKHYTASTGSLPEILEN  
LIERDSGVVYCATWDRITYKKLFGSGTTLVTDKQADAVSPKPTFLPSIAETKQ  
KAGTYLCLEKFPDIIKIHMOKESNTLIGSOEGMTKNDYMFSLTVEESLD  
KEHRCIVHNNKNGIDOEIIFPIKDTVDPKRNVSKDANDVTMDPKDMSKDANDTLQLTNTSAYVWYL  
LLKSVYFAIITCCLGRTRFCNGEKS"

misc\_feature 116..521  
/note="V(g)8-J(g)2 region"  
misc\_feature 158..159  
/note="exon/exon boundary"  
misc\_feature 464..475  
/note="pnt. N-region 5"  
misc\_feature 522..1135  
/note="C(g)2 region"  
misc\_feature 803..808  
/note="pnt. polyA signal"  
misc\_feature 841..846  
/note="pnt. polyA signal"  
misc\_feature 851..852  
/note="exon/exon boundary"  
misc\_feature 899..900  
/note="exon/exon boundary"  
misc\_feature 947..948  
/note="exon/exon boundary"  
misc\_feature 995..996  
/note="exon/exon boundary"

## ORIGIN

Alignment Scores:  
Pred. No.: 5.36e-28 Length: 1160  
Score: 294.00 Matches: 55  
Percent Similarity: 98.28% Conservative: 2  
Best Local Similarity: 94.83% Mismatches: 1  
Query Match: 95.77% Indels: 0  
DB: 9 Gaps: 0

US-10-031-158-14 (1-58) x HSTRGC64 (1-1160)

Qy 1 MetGlnMetPheProPseProLeuPhePhePheLeuGlnLeuLeuLyGlnSerSer 20  
Db 534 ATGCGATGTTTCCCGCAAGCCCACTATTTCTTCCTTCATTCGTAACAAACTCC 593

Qy 21 ArgArgLeuGlnHisIsthPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40  
Db 594 AGAAGCGCTGGAACATACCTTGTCTCTTGAGAAATTTTCCAGATATTATTAAGATAC 653

Qy 41 TleGlyLysLysArgArgAlaThrArgPheTyrAspProArgArgGlyThrPro 58  
Db 654 ATTGGCAAGAAAGAGAGCAACGATTCGGATCCCAAGAGGGAACACCA 707

## RESULT 38

HUMTCR6B 1402 bp mRNA linear PRI 14-JAN-1995  
LOCUS Human T-cell receptor gamma chain VCTI-3(CII)-CIII region mRNA. 3'  
DEFINITION end.  
X06784

ACCESSION M16804.1 GI:339408  
VERSION T-cell receptor  
KEYWORDS T-cell receptor  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1402)  
AUTHORS Krangel, M.S., Band, H., Hata, S., McLean, J. and Brenner, M.B.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

TITLE	Structurally divergent human T cell receptor gamma proteins encoded by distinct C gamma genes							
JOURNAL	Science 237 (4810), 64-67 (1987)							
MEDLINE	8726353							
PUBMED	2955517							
COMMENT	<p>Original source text: Human (cell line IDP2), cDNA to mRNA, clone IDP2.11.</p> <p>Only a few V and J region genes occur within the genome. Diversity in T-cell receptors occur in the V-J recombination events. In some cases, this produces unproductively rearranged reading frames. Also, in some recombination events, some gene regions can be included more than one time creating even more diversity in TCR's.</p> <p>location/Qualifiers</p>							
FEATURES								
source	1..1402 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /map="7p15-p14" 1..1402 /gene="TCRG" <1..1402 /gene="TCRG" /product="TCR-gamma mRNA" <1..957 /gene="TCRG" /note="T-cell receptor (V-J-C) precursor" /codon_start=1 /protein_id="AA61124.1" /db_xref="GI:339409" /db_xref="GDB:G00-120-407" /translation="GGPRISSTKTLSTARLECVSGITTSATSVYWRBPGVIOFLVSISSDYCTAKESGISPSGRKEVDRIPEPTSTLTINHEKODIATYCAMEGYKKLFGSGITLVTDKQLDADVSPFPIELPSIAETLKQKAGTYLCLEKFPIILKIHWBKSNKITLSQSGNTMKNIDNYMEKSMVLTVEEKSLDKSHRCIVRHENKNKDIDEIIIEPLITDVTVDPKDSYSKQANDVTTVDPKVYSKQANDVITMDPKNCSKQANDTLTLLQLTNISAVYTWVLLLLLKSVVFALITCLLRTAFCNGERS"							
CDS	misc_recomb /gene="TCRG" /organism="Homo sapiens" 340..341 /gene="TCRG" /organism="Homo sapiens" misc_recomb /gene="TCRG" /organism="Homo sapiens" 670..671 /gene="TCRG" /organism="Homo sapiens" 716..717 /gene="TCRG" /organism="Homo sapiens" 766..767 /gene="TCRG" /organism="Homo sapiens" 814..815 /gene="TCRG" /organism="Homo sapiens"							
misc_recomb	misc_recomb /gene="TCRG" /organism="Homo sapiens" 340..341 /gene="TCRG" /organism="Homo sapiens"							
misc_recomb	misc_recomb /gene="TCRG" /organism="Homo sapiens" 670..671 /gene="TCRG" /organism="Homo sapiens"							
misc_recomb	misc_recomb /gene="TCRG" /organism="Homo sapiens" 716..717 /gene="TCRG" /organism="Homo sapiens"							
misc_recomb	misc_recomb /gene="TCRG" /organism="Homo sapiens" 766..767 /gene="TCRG" /organism="Homo sapiens"							
misc_recomb	misc_recomb /gene="TCRG" /organism="Homo sapiens" 814..815 /gene="TCRG" /organism="Homo sapiens"							
ORIGIN	Chromosome 7p15.							
Alignment Scores:								
Pred. No.:	6,62e-28	Length:	1402					
Score:	294.00	Matches:	55					
Percent Similarity:	98.28%	Conservative:	2					
Best Local Similarity:	94.83%	Mismatch:	1					
Query Match:	95.77%	Indels:	0					
DB:	9	Gaps:	0					
US-10-031-158-14 (1-58) x HUMTCRGB (1-1402)								
Oy	1	MetGlmetPheProPseSerProleuPhePheLeuGlnLeuleuysGInserSer	20					
Db	353	ATGCAGATGTCCCAACCAAGCCCATATTTCCTCCTCAATGCTGAACAACAATCC	412					
Oy	21	ArgAagLeuGlnHisThrPheValPheLeuAAGAanPheSerLeuMetLeuleuArgTy	40					
Db	413	AGAAAGCTCGAACATACTTGCTGCTTGGAAGAAATTTTCCAGATATTATTAAAGATAC	472					

QY	DB	RESULT 39	LOCUS	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	ORIGIN	Alignment Scores:	Pred. NO.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	DB:
41	1	116131yylvysaargalathnArxPheTyrpProArGAGlyThPro 58	AX332835	Sequence 3344 from Patent WO0194629.	1421 bp	DNA	linear	PAT 09-JAN-2002							6,72e-28	Length: 1421	Matches: 55	Conservative: 2	Mismatches: 1	Indels: 0	Gaps: 0
423	423	ATGGAGATGTTCCCCCAAGCCACCTATTTCTTCCTCAATTCGAAACAAACCTCC 482	AX332835	GI:18123469											294.00	Matched: 55	Conservative: 2	Mismatches: 1	Indels: 0	Gaps: 0	
21	483	ArgargLeugluhsthrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40	AX332835	GI:18123469											98.28%	Matched: 55	Conservative: 2	Mismatches: 1	Indels: 0	Gaps: 0	
483	483	AGAAAGCTCGGAACATACCTTTGTCTCTTGGAATTTTCCAGATATTATTAAGATAC 542	AX332835	GI:18123469											94.83%	Matched: 55	Conservative: 2	Mismatches: 1	Indels: 0	Gaps: 0	
41	41	116131yylvysaargalathnArxPheTyrpProArGAGlyThPro 58	AX332835	GI:18123469											95.77%	Matched: 55	Conservative: 2	Mismatches: 1	Indels: 0	Gaps: 0	
543	543	ATTGGCAAGAAAGAGAGCAACGATTCGTGGATCCAGAGGGAGCAACCA 596	AX332835	GI:18123469																	
1	1	MetGlnMetPhePProPProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20	AX332835	GI:18123469																	
423	423	ATGGAGATGTTCCCCCAAGCCACCTATTTCTTCCTCAATTCGAAACAAACCTCC 482	AX332835	GI:18123469																	
21	21	ArgargLeugluhsthrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40	AX332835	GI:18123469																	
483	483	AGAAAGCTCGGAACATACCTTTGTCTCTTGGAATTTTCCAGATATTATTAAGATAC 542	AX332835	GI:18123469																	
41	41	116131yylvysaargalathnArxPheTyrpProArGAGlyThPro 58	AX332835	GI:18123469																	
543	543	ATTGGCAAGAAAGAGAGCAACGATTCGTGGATCCAGAGGGAGCAACCA 596	AX332835	GI:18123469																	
1	1	MetGlnMetPhePProPProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20	AX332835	GI:18123469																	
423	423	ATGGAGATGTTCCCCCAAGCCACCTATTTCTTCCTCAATTCGAAACAAACCTCC 482	AX332835	GI:18123469																	
21	21	ArgargLeugluhsthrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40	AX332835	GI:18123469																	
483	483	AGAAAGCTCGGAACATACCTTTGTCTCTTGGAATTTTCCAGATATTATTAAGATAC 542	AX332835	GI:18123469																	
41	41	116131yylvysaargalathnArxPheTyrpProArGAGlyThPro 58	AX332835	GI:18123469																	
543	543	ATTGGCAAGAAAGAGAGCAACGATTCGTGGATCCAGAGGGAGCAACCA 596	AX332835	GI:18123469																	
1	1	MetGlnMetPhePProPProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20	AX332835	GI:18123469																	
423	423	ATGGAGATGTTCCCCCAAGCCACCTATTTCTTCCTCAATTCGAAACAAACCTCC 482	AX332835	GI:18123469																	
21	21	ArgargLeugluhsthrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40	AX332835	GI:18123469																	
483	483	AGAAAGCTCGGAACATACCTTTGTCTCTTGGAATTTTCCAGATATTATTAAGATAC 542	AX332835	GI:18123469																	
41	41	116131yylvysaargalathnArxPheTyrpProArGAGlyThPro 58	AX332835	GI:18123469																	
543	543	ATTGGCAAGAAAGAGAGCAACGATTCGTGGATCCAGAGGGAGCAACCA 596	AX332835	GI:18123469																	
1	1	MetGlnMetPhePProPProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20																			

## COMMENT

Original source text: Human T-cell line HPB-MLT, cDNA to mRNA, clones pT-gamma-[1.2].  
Computer-readable sequence of [1] kindly provided by D.P. Dialynas, 31-Oct-1986.  
An aberrant rearrangement at the V-J junction causes a frame shift that leads to peptide termination at the 'taa' codon at positions 370-372.  
A polyadenylation signal can be found at positions 1400 to 1405.

## FEATURES

source  
location/Qualifiers  
1..1421  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
4..979  
/note="pseudo-TCG pept"  
/pseudo  
/codon\_start=1  
4..63  
/note="pseudo-TCG signal pept"  
/pseudo  
64..976  
/note="pseudo-TCG mature pept"  
/pseudo  
mat\_peptide  
16 bp upstream of Apai site; chromosome 7.

## Alignment Scores:

Pred. No.:	6.72e-28	Length:	1421
Score:	294.00	Matches:	55
Percent Similarity:	98.28%	Conservative:	2
Best Local Similarity:	94.83%	Mismatches:	1
Query Match:	95.77%	Indels:	0
DB:	9	Gaps:	0

US-10-031-158-14 (1-58) x HUMTCGNH (1-1421)

Qy	1	MetGlnMetPheProPheSerProLeuPhePhePheLeuGlnLeuLeuGlnSerSer	20
Db	423	ATGCAGATCTTCCCCCAAGCCCACTATTCTTCTTCAATGCTGAAACAAACTCC	482
Qy	21	ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr	40
Db	483	AGAAAGCTGGAACATACCTTGTCTCTTGAGAAATTTTCCAGATATTATTAAAGATAC	542
Qy	41	IleGlyValValArgArgAlaThrArgPheTTPAspProArgArgGlyThrPro	58
Db	543	ATTGGCAAGAAAGAGAGCAACACGATTCTGGATCCCGAGAGGAGAACACCA	596

Search completed: December 8, 2004, 08:54:17  
Job time : 1916 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_p2n model

Run on: December 8, 2004, 06:09:26 ; Search time 247 Seconds  
(without alignments)  
1232.658 Million cell updates/sec

Title: US-10-031-158-14

Perfect score: 307  
Sequence: 1 MQMFPSPDLFFFLQLLKQS.....RYGKKRATRFMDPRGTP 58

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 413486 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

MODEL=frame\_p2n.model -DEV=xlh  
-O=/cgn2.1/USPTO-spool/US1001158/runat\_06122004\_082658\_15877/app\_query.fasta\_1.199  
-DB=N\_Geneseq\_23Sep04 -QFMT=fastap -SUFFIX=ing -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40  
-MODE=LOCAL -OUTFMT=prto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US1001158\_@CGN\_1\_1\_352\_@runat\_06122004\_082658\_15877 -NCPU=6 -ICPU=3  
-NO KWAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOC  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_23Sep04:\*

1: geneseq19908:\*\n2: geneseq19908:\*\n3: geneseq20008:\*\n4: geneseq2001as:\*\n5: geneseq2001bs:\*\n6: geneseq2002as:\*\n7: geneseq2002bs:\*\n8: geneseq2003as:\*\n9: geneseq2003bs:\*\n10: geneseq2003cs:\*\n11: geneseq2003ds:\*\n12: geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	307	100.0	426	3 AAC00681	AAC00681 Human sec
2	307	100.0	470	4 AAK26982	AAK26982 Human bon
3	307	100.0	470	6 ABS01569	ABS01569 Human gen
4	307	100.0	477	4 ABA43588	ABA43588 Human bre
5	307	100.0	477	4 AAK27758	AAK27758 Human bon
6	307	100.0	477	6 ABS02210	ABS02210 Human gen

C	7	307	100.0	539	5	ABV28618	Abv28618 Human pro
C	8	307	100.0	829	2	ABV22791	Abv22791 Human pro
C	9	307	100.0	535	5	AAQ66888	AAQ66888 Human lym
C	10	307	100.0	825	9	ACF35988	ACF35988 Human T c
C	11	307	100.0	1027	4	AAFS6401	AAFS6401 Human TCR
C	12	307	100.0	1027	10	ADB75580	ADB75580 Prostate
C	13	307	100.0	1316	5	ABV26485	Abv26485 Human pro
C	14	307	100.0	1316	5	ABV25093	Abv25093 Human pro
C	15	307	100.0	1395	12	ADL06443	ADL06443 Human tum
C	16	307	100.0	1418	6	AAD38828	AA38828 Human PSN
C	17	307	100.0	2658	10	ADE08798	Ade08798 Novel DNA
C	18	304	99.0	486	5	ABV25013	Abv25013 Human pro
C	19	302	98.4	316	4	ABA47985	ABA47985 Human bre
C	20	302	98.4	316	4	AAK40015	AAK40015 Human bon
C	21	302	98.4	316	6	ABSI4101	ABSI4101 Human gen
C	22	298	97.1	569	5	ABV43475	Abv43475 Human pro
C	23	294	95.8	1421	6	ABL65007	ABL65007 Lung cancr
C	24	294	95.8	1560	10	ADE06987	Ade06987 Novel cod
C	25	294	95.8	1799	6	ABZ35381	ABZ35381 Human gen
C	26	290	94.5	1080	2	AAQ37617	AAQ37617 TCR gamma
C	27	290	94.5	2658	10	ADE08798	Ade08798 Novel DNA
C	28	284	92.5	1046	12	ADP10449	ADP10449 Reference
C	29	284	92.5	1586	6	ABK64529	ABK64529 Human den
C	30	279	90.9	1080	1	AAV91698	AAV91698 Sequence
C	31	259	84.4	533	5	ABV37856	ABV37856 Human pro
C	32	251	81.8	460	3	AAAC04122	AAAC04122 Human sec
C	33	245	79.8	1155	3	AAFI5623	AAFI5623 Human pro
C	34	240	78.2	1022	5	ABV29989	ABV29989 Human pro
C	35	240	78.2	1022	5	ABV24072	ABV24072 Human pro
C	36	240	78.2	1022	5	ABV24124	ABV24124 Human pro
C	37	240	78.2	1022	5	ABV25823	ABV25823 Human pro
C	38	240	78.2	1022	5	ABV24034	ABV24034 Human pro
C	39	240	78.2	1022	5	ABV23986	ABV23986 Human pro
C	40	240	78.2	1022	5	ABV23991	ABV23991 Human pro
C	41	240	78.2	1022	5	ABV24108	ABV24108 Human pro
C	42	240	78.2	1022	5	ABV23814	ABV23814 Human pro
C	43	240	78.2	1022	5	ABV29918	ABV29918 Human pro
C	44	240	78.2	1022	5	ABV29874	ABV29874 Human pro
C	45	240	78.2	1022	5	ABV29953	ABV29953 Human pro

## ALIGNMENTS

RESULT 1	
AAAC00681	standard; CDNA; 426 BP.
AC	AAAC00681;
AC	AAAC00681;
DT	06-OCT-2000 (first entry)
DE	Human secreted protein 5' EST, SEQ ID NO: 679.
DE	Human, 5' EST, expressed sequence tag; secreted protein; CDNA isolation;
KW	gene therapy; chromosome mapping; ss.
KW	
OS	Homo sapiens.
PN	EP1033401-A2.
PD	06-SEP-2000.
PF	21-FEB-2000; 2000EP-00200610.
PR	26-FEB-1999; 99US-0122487P.
PA	(GENSET) GENSET.
PI	Dumas Milne Edwards J, Duclert A, Giordano J;
DR	WPI; 2000-500381/45.
DR	P-PSDB; AAG00675.

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures.

PS Claim 1, SEQ ID NO 679; 71bp + Sequence Listing; English.

XX The present sequence is one of a large number of 5' ESTs derived from  
 CC mRNAs encoding secreted proteins. An ORF has been identified within the  
 CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs  
 CC derived from 30 different tissues. EST sequences usually correspond  
 CC mainly to the 3' untranslated region (UTR) of the mRNA because they are  
 CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not  
 CC well suited for isolating cDNA sequences derived from the 5' ends of  
 CC mRNAs and even in those cases where longer cDNA sequences have been  
 CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from  
 CC mRNAs with intact 5' ends and can therefore be used to obtain full length  
 CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,  
 CC gene therapy and chromosome mapping procedures. They are used to obtain  
 CC upstream regulatory sequences and to design expression and secretion  
 CC vectors

XX SQ Sequence 426 BP; 151 A; 83 C; 88 G; 104 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
4, 87e-34	426	58	0	0	0	0
Score:	307.00	58	0	0	0	0
Percent Similarity:	100.00%	100.00%	0	0	0	0
Best Local Similarity:	100.00%	100.00%	0	0	0	0
Query Match:	3	3	0	0	0	0

US-10-031-158-14 (1-58) x AAC00681 (1-426)

QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSer 20

DB 94 ATGCAAGATGTTCCCAAGCCCACTATTTCTTCCTCAATGCTGAAACAAAGCTCC 153

QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40

DB 154 AGAAGGCTGGACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAATAC 213

QY 41 TleGlyLysLysArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58

DB 214 ATTGGCAAGAAAGAGAGACACATCTGGGATCCCGAGGAGGAAACACCA 267

#### RESULT 2

ID AAK26982 standard; DNA; 470 BP.

XX AAK26982;

DT 06-NOV-2001 (first entry)

DE Human bone marrow expressed single exon probe SEQ ID NO: 1539.

KM Human; bone marrow expressed exon; gene expression analysis; probe;

KM microarray; cancer; leukemia; lymphoma; myeloma; ss.

OS Homo sapiens.

PN WO200157276-A2.

XX 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000668.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human bone marrow.

PS Example 4; SEQ ID NO 1539; 658bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC bone marrow. They can be used to measure gene expression in bone marrow  
 CC samples, which may enable the improved diagnosis and treatment of cancers  
 CC such as lymphoma, leukemia and myeloma. The present sequence is one of  
 CC the probes of the invention

XX SQ Sequence 470 BP; 148 A; 102 C; 87 G; 133 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
5, 56e-34	470	58	0	0	0	0
Score:	307.00	58	0	0	0	0
Percent Similarity:	100.00%	100.00%	0	0	0	0
Best Local Similarity:	100.00%	100.00%	0	0	0	0
Query Match:	4	4	0	0	0	0

US-10-031-158-14 (1-58) x AAK26982 (1-470)

QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSer 20

DB 203 ATGCAAGATGTTCCCAAGCCCACTATTTCTTCCTCAATGCTGAAACAAAGCTCC 262

QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40

DB 263 AGAAGGCTGGACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAATAC 322

QY 41 TleGlyLysLysArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58

DB 323 ATTGGCAAGAAAGAGAGACACATCTGGGATCCCGAGGAGGAAACACCA 376

#### RESULT 3

ID ABS01569 standard; DNA; 470 BP.

XX ABS01569;

DT 19-AUG-2002 (first entry)

DE Human genome-derived single exon probe from lung SEQ ID NO 1560.

KM Human; de; single exon probe; asthma; lung cancer; COPD; ILD;

KM chronic obstructive pulmonary disease; interstitial lung disease;

KM familial idiopathic pulmonary fibrosis; neurofibromatosis;

KM tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;

KM Heremansky-Pudlak syndrome; sarcoidosis; pulmonary haemorrhage;

KM pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;

KM pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;

KM primary ciliary dyskinesia; pulmonary hypertension;

XX hyaline membrane disease.

OS Homo sapiens.

PN WO200186003-A2.

XX 15-NOV-2001.

PF 30-JAN-2001; 2001WO-US000665.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2002-114183/15.  
 DR  
 PT Spatially-addressable set of single exon nucleic acid probes, used to  
 PT measure gene expression in human lung samples.  
 XX  
 PS Claim 1; SEQ ID NO 1560; 634pp; English.  
 CC  
 CC The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human lung comprising single exon nucleic acid probes having one of  
 CC 12614 nucleic acid sequences mentioned in the specification, or their  
 CC complements or the 12387 open reading frames derived from the 12614  
 CC probes. Also included are a microarray comprising the novel set of probes  
 CC; the novel set of probes which hybridize at high stringency to a nucleic  
 CC acid expressed in the human lung; measuring gene expression in a sample  
 CC derived from human lung, comprising (a) contacting the array with a  
 CC collection of detectably labeled nucleic acids derived from human lung  
 CC mRNA, and (b) measuring the label detectably bound to each probe of the  
 CC array; identifying exons in a eukaryotic genome, comprising (a)  
 CC algorithmically predicting at least one exon from genomic sequences of  
 CC the eukaryote; and (b) detecting specific hybridisation of detectably  
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
 CC having a fragment identical to the predicted exon, the probe is included  
 CC in the above mentioned microarray; assigning exons to a single gene,  
 CC comprising (a) identifying exons from genomic sequence by the method  
 CC above and (b) measuring the expression of each of the exons in several  
 CC tissues and/or cell types using hybridisation to a single exon  
 CC microarray having a probe with the exon, where a common pattern of  
 CC expression of the exons in the tissues and/or cell types indicates that  
 CC the exons should be assigned to a single gene; a peptide comprising one  
 CC of 12011 sequences, mentioned in the specification, or encoded by the  
 CC probes/open reading frames (ORF). The probes are used for gene expression  
 CC analysis, and for identifying exons in a gene, particularly using human  
 CC lung derived mRNA and for the study of lung diseases such as asthma, lung  
 CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung  
 CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,  
 CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-  
 CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary  
 CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,  
 CC Kargener syndrome, fibrocystic pulmonary dysplasia, primary ciliary  
 CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The  
 CC present sequence is a single exon probe of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX SQ Sequence 470 BP; 148 A; 102 C; 87 G; 133 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 5,56e-34 Length: 470  
 Score: 307.00 Matches: 58  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: Gaps: 0  
 US-10-031-158-14 (1-58) x ABS01569 (1-470)  
 QY 1 MetGIMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20  
 DB 203 ATGCAAGATGTTCCGCCAAGCCACATATTTCTTCCTCAATGCTGAACAAAGCTCC 262  
 QY 21 AAGAGTLeuGlnHisThrPheValPheLeuArgAsnPhaSerLeuMetLeuLeuArgTyr 40

DB 263 AGAAGGCTGAACATACCTTGTCTTGTGAGAAATTTTCCGTATGTATTAAGATAC 322  
 QY 41 TleGlyLysLysAAGAGTAAATThrArgPheThrPheProAAGAGGlyThPro 58  
 DB 323 ATTGGCAAGAAAGAGAGCAACAGATTCTGGATCCAGAGGGGAACACCA 376  
 RESULT 4  
 ABA43588  
 ID ABA43588 standard; DNA; 477 BP.  
 XX  
 XX AC ABA43588;  
 XX  
 XX DT 01-FEB-2002 (first entry)  
 XX  
 DE Human breast cell single exon nucleic acid probe #2283.  
 KW Human; microarray; single exon probe; gene expression; breast; disease;  
 KW cancer; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157271-A2.  
 PD  
 PD 09-AUG-2001.  
 XX  
 PP 30-JAN-2001; 2001WO-US000662.  
 XX  
 PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207455P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-496933/54.  
 DR  
 PT New spatially-addressable set of single exon nucleic acid probes, useful  
 PT for measuring gene expression in sample derived from human breast, useful  
 PT comprises number of single exon nucleic acid probes.  
 XX  
 PS Claim 1; SEQ ID NO 2283; 327pp + Sequence Listing; English.  
 XX  
 CC The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human breast and BT 474 cells. The method involves contacting the  
 CC probes with a collection of detectably labelled nucleic acids derived  
 CC from mRNA of human breast, and then measuring the label bound to each  
 CC probe of the microarray. The probes are useful for verifying the  
 CC expression of regions of genomic DNA predicted to encode proteins. They  
 CC are useful for gene discovery, and for determining predisposition and/or  
 CC prognosing breast disease. Gene expression analysis is useful for  
 CC assessing the toxicity of chemical agents on cells. The microarray of  
 CC this invention presents a far greater diversity of probes for measuring  
 CC gene expression, with far less bias than expressed sequence tag  
 CC microarrays. The method is suitable for rapid production of functional  
 CC information from genomic sequence. The present sequence is a single exon  
 CC nucleic acid probe of the invention. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX SQ Sequence 477 BP; 147 A; 104 C; 88 G; 138 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 5,67e-34 Length: 477  
 Score: 307.00 Matches: 58  
 Percent Similarity: 100.00% Conservative: 0

```

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0
US-10-031-158-14 (1-58) x ABA43588 (1-477)

Oy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSer 20
    |||
Db 212 ATGCAGATGTTTCCCCCAAGCCCACTATTCTTCTCAATTGCTGAACAAAGCTCC 271
    |||

Oy 21 ArgATGLeuGluHisThrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40
    |||
Db 272 AGAAGGCTGGAACATACCTTTGCTTGAAGAAATTTTCCCTGATGTATTAAAGATAC 331
    |||

Oy 41 IlegLYsLYsArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58
    |||
Db 332 ATTGGCAAGAAAGAAAGAGACACACATTTCTGGGATCCCAAGAGGGAACACCA 385
    |||

RESULT 5
AAK27758
ID AAK27758 standard; DNA; 477 BP.
XX
AC AAK27758;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 2315.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KM microarray; cancer; leukemia; lymphoma; myeloma; ss.
OS Homo sapiens.
XX
PN W0200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000668.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human bone marrow.
XX
PS Example 4; SEQ ID NO 2315; 658bp + Sequence listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to measure gene expression in bone marrow
XX samples, which may enable the improved diagnosis and treatment of cancers
XX such as lymphoma, leukemia and myeloma. The present sequence is one of
XX the probes of the invention
XX
SQ Sequence 477 BP; 147 A; 104 C; 88 G; 138 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5,67e-34 Length: 477
Score: 307.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0

```

```

DB: 4 Gaps: 0
US-10-031-158-14 (1-58) x AAK27758 (1-477)

Oy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSer 20
    |||
Db 212 ATGCAGATGTTTCCCCCAAGCCCACTATTCTTCTCAATTGCTGAACAAAGCTCC 271
    |||

Oy 21 ArgATGLeuGluHisThrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40
    |||
Db 272 AGAAGGCTGGAACATACCTTTGCTTGAAGAAATTTTCCCTGATGTATTAAAGATAC 331
    |||

Oy 41 IlegLYsLYsArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58
    |||
Db 332 ATTGGCAAGAAAGAAAGAGACACACATTTCTGGGATCCCAAGAGGGAACACCA 385
    |||

RESULT 6
ABS02210
ID ABS02210 standard; DNA; 477 BP.
XX
AC ABS02210;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human genome-derived single exon probe from lung SEQ ID NO 2201.
XX
KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KM chronic obstructive pulmonary disease; interstitial lung disease;
KM familial idiopathic pulmonary fibrosis; neurofibromatosis;
KM tuberous sclerosis; Gaucher's disease; Niemann-pick disease;
KM Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KM pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;
KM pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KM primary ciliary dyskinesia; pulmonary hypertension;
KM hyaline membrane disease.
XX
OS Homo sapiens.
XX
PN W0200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US000665.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2002-114183/15.
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
XX measure gene expression in human lung samples.
XX
PS Claim 1; SEQ ID NO 2201; 634bp; English.
XX
CC The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human lung comprising single exon nucleic acid probes having one of
XX 12614 nucleic acid sequences mentioned in the specification, or their
XX complements or the 12617 open reading frames derived from the 12614
XX probes. Also included are a microarray comprising the novel set of probes
XX acid expressed in the human lung; measuring gene expression in a sample
XX derived from human lung, comprising (a) contacting the array with a
XX collection of detectably labeled nucleic acids derived from human lung

```

CC mRNA, and (b) measuring the label detectably bound to each probe of the  
CC array; identifying exons in a eukaryotic genome, comprising (a)  
CC algorithmically predicting at least one exon from genomic sequences of  
CC the eukaryote; and (b) detecting specific hybridisation of detectably  
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
CC having a fragment identical to the predicted exon, the probe is included  
CC in the above mentioned microarray; assigning exons to a single gene,  
CC comprising (a) identifying exons from genomic sequence by the method  
CC above and (b) measuring the expression of each of the exons in several  
CC tissues and/or cell types using hybridisation to a single exon  
CC microarrays having a probe with the exon, where a common pattern of  
CC expression of the exons in the tissues and/or cell types indicates that  
CC the exons should be assigned to a single gene; a peptide comprising one  
CC of 1201 sequences, mentioned in the specification, or encoded by the  
CC probes/open reading frames (ORF). The probes are used for gene expression  
CC analysis, and for identifying exons in a gene, particularly using human  
CC lung derived mRNA and for the study of lung diseases such as asthma, lung  
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung  
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,  
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Heremakly-  
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary  
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,  
CC Karsenger syndrome, fibrocystic pulmonary dysplasia, primary ciliary  
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The  
CC present sequence is a single exon probe of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences

CC SQ Sequence 477 BP; 147 A; 104 C; 88 G; 138 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5,67e-34 Length: 477  
Score: 307.00 Matches: 58  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-031-158-14 (1-58) x ABS02210 (1-477)

QY 1 MetGlnMetPheProSerProLeuPhePhePheLeuGlnLeuLeuysGlnSerSer 20  
DB 212 ATGCAGATGTTTCCCAAGCCCACTATTTTCTTCAATGCTGAAACAAAGCTCC 271  
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40  
DB 272 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAAGATAC 331  
QY 41 TleGlyLysLysArgArgAlaThrArgPheTTPAAPPProArgArgGlyThrPro 58  
DB 332 ATTGGCAAGAAAGAAAGAGCAACGATTTCTGGATCCAGAGGGGAACACCA 385

RESULT 7

ABV28618/C  
ID ABV28618 standard; cDNA, 539 BP.

XX AC ABV28618;  
XX DT 16-SEP-2002 (first entry)  
XX DE Human prostate expression marker cDNA 28609.  
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
XX KW pharmacogenomic marker; gene; ss.  
XX OS Homo sapiens.  
XX XX  
XX PN WO200160860-A2.  
XX PD 23-AUG-2001.  
XX PF 20-FEB-2001; 2001WO-US005171.

XX 17-FEB-2000; 2000US-0183131P.  
PR 16-MAR-2000; 2000US-0189863P.  
PR 25-MAY-2000; 2000US-0207454P.  
PR 09-JUN-2000; 2000US-0211314P.  
PR 18-JUL-2000; 2000US-0219007P.  
PR 13-DEC-2000; 2000US-0255281P.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Endage WO, Monahan JE;

DR WPI; 2001-662795/76.

PT Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1; Page 5996; 11750p; English.

CC The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for: (a) assessing whether  
CC a patient is afflicted with prostate cancer; (b) monitoring the  
CC progression of prostate cancer in a patient; (c) assessing the efficacy  
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
CC determining whether prostate cancer has metastasized in a patient; (h)  
CC assessing the aggressiveness or indolence of prostate cancer in a patient  
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker

XX SQ Sequence 539 BP; 134 A; 104 C; 106 G; 195 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 6,68e-34 Length: 539  
Score: 307.00 Matches: 58  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 5 Gaps: 0

US-10-031-158-14 (1-58) x ABV28618 (1-539)

QY 1 MetGlnMetPheProSerProLeuPhePhePheLeuGlnLeuLeuysGlnSerSer 20  
DB 410 ATGCAGATGTTTCCCAAGCCCACTATTTTCTTCAATGCTGAAACAAAGCTCC 351  
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40  
DB 350 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAAGATAC 291  
QY 41 TleGlyLysLysArgArgAlaThrArgPheTTPAAPPProArgArgGlyThrPro 58  
DB 290 ATTGGCAAGAAAGAAAGAGCAACGATTTCTGGATCCAGAGGGGAACACCA 237

RESULT 8

ABV22791/C  
ID ABV22791 standard; cDNA, 539 BP.

XX AC ABV22791;  
XX DT 13-SEP-2002 (first entry)  
XX DE Human prostate expression marker cDNA 22782.  
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
XX KW pharmacogenomic marker; gene; ss.  
XX OS Homo sapiens.  
XX XX  
XX PN WO200160860-A2.

```

XX 23-AUG-2001.
PD
XX
XX 20-FEB-2001; 2001WO-US005171.
PF
XX
XX 17-FEB-2000; 2000US-0183119P.
PR 16-MAR-2000; 2000US-0189662P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan UE;
XX
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 4022; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (1) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement; (1) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or incidence of prostate cancer in a patient
CC / (1) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX Sequence 539 BP; 134 A; 104 C; 106 G; 195 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 6.68e-34 Length: 539
Score: 307.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-10-031-158-14 (1-58) x ABV22791 (1-539)
QY 1 MetGlnMetPheProPheProSerProLeuPhePhePheLeuGlnLeuLeuYsgInSerSer 20
Db 410 ATGCAGATGTTTCCCCCAAGCCCACTATTTTCTTCTTCAATTCGAAACAAAGCTCC 351
QY 21 ArgArgLeuGlnIuhIsthrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40
Db 350 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAAGTAC 291
QY 41 IleGlyLysLysArgArgAlaThrArgPheThrAspProArgArgGlyThrPro 58
Db 290 ATTGGCAAGAAAGAAAGAGCAACAGATTCGGATCCAGAGGAGAAACCA 237
RESULT 9
AAQ66888
ID AAQ66888 standard; cDNA; 825 BP.
XX
XX AAQ66888;
XX
XX 25-MAR-2003 (revised)
DT 02-FEB-1995 (first entry)
XX
XX Human lymphocyte clone G 115 soluble TCR Vgamma chain.
DE
XX Soluble T cell receptor; TCR; pathological T lymphocyte; PCR primer;
KM

```

```

KW polymerase chain reaction; leukaemia; Burkitt's lymphoma;
KW autoimmune disease; cell typing; gamma; delta; lymphocyte G115;
KW Mycobacterium tuberculosis; ss.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH CDS 1..825
FT /*tag= a
FT /product= "soluble_Vgamma_chain"
XX
XX WO9412648-A2.
XX
XX 09-JUN-1994.
XX
XX 25-NOV-1993; 93WO-FR001165.
XX
XX 25-NOV-1992; 92FR-00014203.
XX
XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
PA (IMMU-) IMMUNOTECH SA.
XX
XX Bonnevillie M;
PI
XX
XX WPI; 1994-200276/24.
XX
XX P-PSDB; AAR5705.
XX
XX Prodn. of soluble T receptors lacking the trans-membrane region - by co-
PT transfection of host cells with truncated sub unit DNA, for diagnosis and
PT treatment of pathological T cell proliferation.
XX
XX Example 1; Fig 2B; 41pp; French.
XX
XX The human lymphocyte clone gamma-delta-s G115 expresses T cell receptors
CC V9J9P1gama/V2D3J1delta. The soluble delta and gamma chains have the
CC sequences AAQ66887 and AAQ66888, respectively. This clone, which kills
CC Burkitt's lymphoma ("Dad1") cells and recognises an antigen present in
CC water-soluble extracts of M.tuberculosis, was used for the construction
CC of gamma-delta-s genes and expression of soluble T cell receptors. The
CC soluble TCRs (i.e. lacking the transmembrane region) are useful in
CC diagnosis (esp. cell typing) and treatment of conditions involving T cell
CC proliferation, e.g. leukaemia, lymphoma or some autoimmune diseases.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 825 BP; 267 A; 178 C; 176 G; 204 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 1.19e-33 Length: 825
Score: 307.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-10-031-158-14 (1-58) x AAQ66888 (1-825)
QY 1 MetGlnMetPheProPheProSerProLeuPhePhePheLeuGlnLeuLeuYsgInSerSer 20
Db 440 ATGCAGATGTTTCCCCCAAGCCCACTATTTTCTTCTTCAATTCGAAACAAAGCTCC 499
QY 21 ArgArgLeuGlnIuhIsthrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40
Db 500 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAAGTAC 559
QY 41 IleGlyLysLysArgArgAlaThrArgPheThrAspProArgArgGlyThrPro 58
Db 560 ATTGGCAAGAAAGAAAGAGCAACAGATTCGGATCCAGAGGAGAAACCA 613
RESULT 10
ACF35988
ID ACF35988 standard; cDNA; 825 BP.
XX
XX ACF35988;
XX

```

```
XX 20-NOV-2003 (first entry)
DT
XX Human T cell receptor Vgamma9 cDNA sequence.
DB
XX T cell receptor; TCR; gastrointestinal; antiinflammatory; cytostatic;
KW tuberculostatic; dermatological; antibacterial; virucide; gynaecological;
KW cell therapy; Vgamma; human; gene; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..825
XX FT /*tag= a
XX FT /product= "TCR Vgamma9"
XX
XX WO2003060097-A2.
XX
XX 24-JUL-2003.
XX
XX 10-JAN-2003; 2003WO-US000728.
XX
XX 10-JAN-2002; 2002US-0347285P.
XX
XX (NAJE-) NAT JEWISH MEDICAL & RES CENT.
XX
XX O'Brien RL, Born WK, Roark CL, Aydinug MK;
XX WPI; 2003-598525/56.
XX P-PSDB; ABR82494.
XX
XX Regular a gammadelta T-cell mediated immune response in a mammal,
XX useful for treating inflammation in intestine, skin, lungs or
XX reproductive tract, comprises administering to the mammal a soluble
XX gammadelta T cell receptor.
XX
XX Example; Page 64-65; 71pp; English.
XX
XX The invention relates to regulating a gammadelta T-cell mediated immune
XX response in a mammal and involves administering to the mammal a soluble
XX gammadelta T cell receptor (TCR). The method is useful for treating
XX patients having, or are at risk of developing an intestinal condition,
XX e.g. Crohn's disease, ischaemic colitis, irritable bowel disease, and
XX colon cancer; a lung condition associated with inflammation such as
XX atway hyperesponsiveness, pneumonia, tuberculosis, and a primary or
XX metastatic lung tumour; a skin condition associated with inflammation
XX such as skin lesion caused by bacterial infection, viral infection or
XX laceration, and a skin cancer; or a condition associated with
XX inflammation of the reproductive tract such as infection caused by
XX bacterial or viral infection that involve the epithelial mucosal lining,
XX a tubal infection, preventing tubal factor infertility, and a cancer
XX selected from ovarian, cervical, uterine, prostate or testicular cancers.
XX Sequences ACF35987-988 represent human TCR Vgamma chain CDNA's used in the
XX present invention
XX
XX Sequence 825 BP; 267 A; 178 C; 176 G; 204 T; 0 U; 0 Other;
SQ
XX
XX Alignment Scores:
XX Pred. No.: 1.19e-33 Length: 825
XX Score: 307.00 Matches: 58
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX Gaps: 0
XX
XX US-10-031-158-14 (1-58) x ACF35988 (1-825)
QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20
DB 440 ATGCAGATGTTTCCCAAGCCCACTATTCTTCAATTCGTAACAAAGCTCC 499
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAnPheSerLeuMetLeuArgTyr 40
|||||
```

```
DB 500 AGAAGGCTGAACATACCTTGTCTTTCGAGAAATTTTCCGTGATGTTATTAAGATAC 559
QY 41 |||GlyLysLysArgArgAlaThrArgPheThrPheProArgArgGlyThrPro 58
DB 560 ATTGCGAAGAAAGAAAGAGCAACAGATTCTGGAGTCCGAGGGAACACCA 613
XX
XX RESULT 11
XX AAF56401
XX ID AAF56401 standard; cDNA, 1027 BP.
XX
XX AC AAF56401;
XX
XX 12-APR-2001 (first entry)
DT
XX
XX Human TCRgamma alternate reading frame protein coding sequence.
DB
XX Human; TARP; prostate cancer; breast cancer; immunotherapy;
KW T cell receptor gamma alternate reading frame protein; TCRgamma; ss.
XX
XX Homo sapiens.
XX
XX WO200104309-A1.
XX
XX 18-JAN-2001.
XX
XX 12-JUL-2000; 2000WO-US019039.
XX
XX 13-JUL-1999; 99US-0143560P.
XX 01-OCT-1999; 99US-0157471P.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Pastan I, Eesand M, Lee B, Vaemacris G, Wolfgang C;
XX WPI; 2001-081050/09.
XX P-PSDB; AAB66399, AAB66400.
XX
XX Isolated T-cell receptor gamma alternate reading frame protein useful for
XX diagnosing and raising an immune response to prostate cancer and breast
XX cancer.
XX
XX Example 1; Fig 1; 85pp; English.
XX
XX The present invention provides the protein and coding sequences of the
XX human T cell receptor alternate reading frame protein (TKRP). This
XX protein is expressed in prostate and breast cancer cells at higher levels
XX than normal and so can be used in the immunotherapy of these cancers, as
XX well as their detection and prevention
XX
XX Sequence 1027 BP; 307 A; 232 C; 200 G; 288 T; 0 U; 0 Other;
SQ
XX
XX Alignment Scores:
XX Pred. No.: 1.59e-33 Length: 1027
XX Score: 307.00 Matches: 58
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX Gaps: 0
XX
XX US-10-031-158-14 (1-58) x AAF56401 (1-1027)
QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20
DB 74 ATGCAGATGTTTCCCAAGCCCACTATTCTTCAATTCGTAACAAAGCTCC 133
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAnPheSerLeuMetLeuArgTyr 40
DB 134 AGAAGGCTGAACATACCTTGTCTTTCGAGAAATTTTCCGTGATGTTAAGATAC 193
QY 41 |||GlyLysLysArgArgAlaThrArgPheThrPheProArgArgGlyThrPro 58
DB 194 ATTGCGAAGAAAGAAAGAGCAACAGATTCTGGAGTCCGAGGGAACACCA 247
|||||
```

RESULT 12  
 ADB75580  
 ID ADB75580 standard; cDNA; 1027 BP.  
 XX  
 AC ADB75580;  
 XX  
 DT 04-DEC-2003 (first entry)  
 XX  
 DE Prostate cancer marker cDNA.  
 XX  
 KM Prostate; cancer; cytostatic; gene therapy; marker; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003009814-A2.  
 XX  
 PD 06-FEB-2003.  
 XX  
 PF 25-JUL-2002; 2002WO-US023913.  
 XX  
 PR 25-JUL-2001; 2001US-0307982P.  
 PR 22-AUG-2001; 2001US-0314356P.  
 PR 25-SEP-2001; 2001US-0325020P.  
 PR 12-DEC-2001; 2001US-0341746P.  
 PR 05-MAR-2002; 2002US-0362158P.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Schlegel R, Monahan JE, Endege WO, Gannavarapu M, Gorbatcheva B;  
 PI Hoersth S, Kamatkar S, Monsey AM, Glatk K, Zhao X, Anderson D;  
 DR WPI; 2003-248033/24.  
 XX  
 PT New nucleic acid molecule, useful for diagnosing or treating prostate  
 cancer.  
 XX  
 PS Disclosure; SEQ ID NO 404; 99pp; English.  
 XX  
 CC The invention relates to newly discovered cancer markers associated with  
 CC the cancerous state of prostate cells. Also disclosed is a method of  
 CC assessing whether a patient is afflicted with prostate cancer. The method  
 CC of the invention involves assessing whether a patient is afflicted with  
 CC prostate cancer by comparing the level of expression of a marker in a  
 CC patient sample and the normal level of expression of the marker in a  
 CC control non-prostate cancer sample, where a significant increase in the  
 CC level of expression of the marker in the patient sample and the normal  
 CC level indicates that the patient is afflicted with prostate cancer.  
 CC Nucleic acids of the invention are useful for diagnosing or treating  
 CC prostate cancer, and may be useful in gene therapy. Sequences given in  
 CC ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence  
 CC data for this patent did not form part of the printed specification, but  
 CC was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 1027 BP; 307 A; 232 C; 200 G; 288 T; 0 U; 0 Other;  
 XX

Alignment Scores:  
 Pred. No.: 1,59e-33 Length: 1027  
 Score: 307.00 Matches: 58  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 10 Gaps: 0

US-10-031-158-14 (1-58) x ADB75580 (1-1027)

QY 1 MetGlnMetPheProSerProLeuPhePhePheLeuGlnLeuLeuysGlnSerSer 20  
 DB 74 ATCGAAGTGTTCCTCCCAAGCCACATATTTCTTCCTTCAATTCGGAACAAAGCTCC 133

QY 21 ArgArgLeuGlnLuhIsthrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTYR 40  
 DB 134 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGATTATTAAGATAC 193

QY 41 IleGlyLysLysArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58  
 DB 194 ATTGGCAAGAAAAGAAAGAGCAACAGATTCTGGATCCACGAGGAGCAACACCA 247

RESULT 13  
 ABV26485/c  
 ID ABV26485 standard; cDNA; 1316 BP.  
 XX  
 AC ABV26485;  
 XX  
 DT 16-SEP-2002 (first entry)  
 XX  
 DE Human prostate expression marker cDNA 26476.  
 XX  
 KM Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
 KM pharmacogenomic marker; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200160860-A2.  
 XX  
 PD 23-AUG-2001.  
 XX  
 PF 20-FEB-2001; 2001WO-US005171.  
 XX  
 PR 17-FEB-2000; 2000US-0189319P.  
 PR 16-MAR-2000; 2000US-0189862P.  
 PR 25-MAY-2000; 2000US-0207454P.  
 PR 09-JUN-2000; 2000US-0211314P.  
 PR 18-JUL-2000; 2000US-0219007P.  
 PR 13-DEC-2000; 2000US-0255281P.  
 XX  
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 XX  
 PI Schlegel R, Endege WO, Monahan JE;  
 PI WPI; 2001-662795/76.  
 DR  
 XX  
 PT Novel isolated nucleic acid molecule associated with cancerous state of  
 PT prostate cells and correlating with presence of prostate cancer, useful  
 PT for detecting presence of prostate cancer, stage of prostate cancer.  
 XX  
 PS Claim 1; Page 5351; 11750pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid molecule (I) comprising  
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
 CC specification or its complement. (I) is useful for: (a) assessing whether  
 CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy  
 CC of progression of prostate cancer in a patient; (c) assessing the efficacy  
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
 CC determining whether prostate cancer has metastasized in a patient; (h)  
 CC assessing the aggressiveness or indolence of prostate cancer in a patient  
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
 XX

Alignment Scores:  
 Pred. No.: 2,22e-33 Length: 1316  
 Score: 307.00 Matches: 58  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 5 Gaps: 0

US-10-031-158-14 (1-58) x ABV26485 (1-1316)

QY 1 MetGlnMetPheProSerProLeuPhePhePheLeuGlnLeuLeuysGlnSerSer 20  
 DB 562 ATCGAAGTGTTCCTCCCAAGCCACATATTTCTTCCTTCAATTCGGAACAAAGCTCC 503

Qy 21 ArgArgLeuGluHisThrPheValPheLeuArgAspPheSerLeuMetLeuLeuArgTyr 40  
 Db 502 AGAAGGCTGGAACATACCTTCTCTTGGAGAAATTTTCCCTGATGTATTAAGATAC 443

Qy 41 IleGlyIysLysArgArgAlaThrArgPheTyrPaspProArgArgGlyThrPro 58  
 Db 442 ATTGGCAAGAAAGAGAGACAGATTCTGGATCCAGAGAGGGAACACCA 389

RESULT 14  
 ABV25093/c  
 ID ABV25093 standard; cDNA; 1316 BP.  
 XX  
 AC ABV25093;  
 XX  
 DT 16-SEP-2002 (first entry)  
 XX  
 DE Human prostate expression marker CDNA 25084.  
 XX  
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
 KW pharmacogenomic marker; gene; ss.  
 OS Homo sapiens.  
 XX  
 PN WO200160860-A2.  
 XX  
 PD 23-AUG-2001.  
 XX  
 PF 20-FEB-2001; 2001WO-US005171.  
 XX  
 PR 17-FEB-2000; 2000US-0183319P.  
 PR 16-MAR-2000; 2000US-0189862P.  
 PR 25-MAY-2000; 2000US-0207454P.  
 PR 09-JUN-2000; 2000US-0211314P.  
 PR 18-JUL-2000; 2000US-021907P.  
 PR 13-DEC-2000; 2000US-0255281P.  
 XX  
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 XX  
 PI Schlegel R, Endege WO, Monahan JE;  
 XX  
 DR WPI; 2001-662795/76.  
 XX  
 PT Novel isolated nucleic acid molecule associated with cancerous state of  
 PT prostate cells and correlating with presence of prostate cancer, useful  
 PT for detecting presence of prostate cancer, stage of prostate cancer.  
 XX  
 PS Claim 1; Page 4881; 11750pp; English.

CC The invention relates to an isolated nucleic acid molecule (I) comprising  
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
 CC specification or its complement. (I) is useful for: (a) assessing whether  
 CC a patient is afflicted with prostate cancer; (b) monitoring the  
 CC progression of prostate cancer in a patient; (c) assessing the efficacy  
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
 CC determining whether prostate cancer has metastasized in a patient; (h)  
 CC assessing the aggressiveness or indolence of prostate cancer in a patient  
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
 XX  
 SQ Sequence 1316 BP; 369 A; 282 C; 300 G; 365 T; 0 U; 0 Other;  
 XX

Alignment Scores:  
 Pred. No.: 2, 22e-33 Length: 1316  
 Score: 307.00 Matches: 58  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 5 Gaps: 0  
 US-10-031-158-14 (1-58) x ABV25093 (1-1316)

Qy 1 MetGlnMetPheProProSerProLeuPhePheGlnLeuLysGlnSerSer 20  
 Db 562 ATGCAGATGTTTCCCGCAAGCCCACTAATTTTCTTCCCTGCAATGCTAAACAAAGCTCC 503

Qy 21 ArgArgLeuGluHisThrPheValPheLeuArgAspPheSerLeuMetLeuLeuArgTyr 40  
 Db 502 AGAAGGCTGGAACATACCTTCTCTTGGAGAAATTTTCCCTGATGTATTAAGATAC 443

Qy 41 IleGlyIysLysArgArgAlaThrArgPheTyrPaspProArgArgGlyThrPro 58  
 Db 442 ATTGGCAAGAAAGAGAGACAGATTCTGGATCCAGAGAGGGAACACCA 389

RESULT 15  
 ADL06443  
 ID ADL06443 standard; cDNA; 1395 BP.  
 XX  
 AC ADL06443;  
 XX  
 DT 20-MAY-2004 (first entry)  
 XX  
 DE Human tumour-associated antigenic target (TAT) cDNA sequence #23.  
 XX  
 KW Human; tumour-associated antigenic target; TAT; cell death; tumour;  
 KW cancer; cytostatic; gene; ss.  
 OS Homo sapiens.  
 XX  
 PN WO2004016225-A2.  
 XX  
 PD 26-FEB-2004.  
 XX  
 PF 19-AUG-2003; 2003WO-US025892.  
 XX  
 PR 19-AUG-2002; 2002US-0404809P.  
 PR 21-AUG-2002; 2002US-0405645P.  
 PR 23-SEP-2002; 2002US-0413192P.  
 PR 15-OCT-2002; 2002US-0419008P.  
 PR 15-NOV-2002; 2002US-0426847P.  
 PR 02-JUL-2003; 2003US-0484959P.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Desauvage PJ, Frantz G, Hillan KJ, Polakis P, Polson A, Smith V;  
 PI Spencer SD, Wu TD, Zhang Z;  
 XX  
 DR WPI; 2004-257144/24.  
 DR P-PsDB; ADL06523.

CC New antibody that binds to a tumor-associated antigenic target (TAT)  
 CC polypeptide, useful for preparing a composition for diagnosing or  
 CC treating cancer.  
 XX  
 PS Claim 1; SEQ ID NO 23; 319pp; English.

CC The present invention relates to the isolation of human tumour-associated  
 CC antigenic target (TAT) polynucleotide and polypeptide sequences. Also  
 CC disclosed is an antibody that binds to a TAT polypeptide. The antibody is  
 CC a monoclonal antibody, an antibody fragment, a chimeric antibody or a  
 CC humanized antibody. It is conjugated to a growth inhibitory agent. It is  
 CC produced in bacteria or in CHO cells and induces death of a cell to which  
 CC it binds. The antibody is useful for preparing a composition for  
 CC diagnosing or treating tumours and cancer. The present sequence  
 CC represents a human TAT cDNA sequence of the invention.  
 XX  
 SQ Sequence 1395 BP; 405 A; 324 C; 286 G; 380 T; 0 U; 0 Other;  
 XX

Alignment Scores:  
 Pred. No.: 2, 41e-33 Length: 1395  
 Score: 307.00 Matches: 58  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0

```

DB: 12 Gaps: 0
US-10-031-158-14 (1-58) x ADL06443 (1-1395)
QY 1 MetGlnMetPheProPheSerProLeuPhePhePheLeuGlnLeuLeuYagInSerSer 20
   |||
DB 449 ATGCAGATGTTTCCCAAGCCACGATTTTCTTCTTCATGTCGAAACAAAGCTCC 508
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40
   |||
DB 509 AGAAGGCTGGACATACCTTTGCTCTTGAGAAATTTTCCCTGATGTATTAAATAC 568
QY 41 IleglyLysLysArgArgAlaThrArgPheTTPaPProArgArgGlyThrPro 58
   |||
DB 569 ATTGGCAAGAAAAGAGACCAACGATTCGGGATCCAGAGGGGAACACCA 622

RESULT 16
AD38828
ID AAD38828 standard; cDNA; 1418 BP.
XX
AC AAD38828;
XX
DT 23-SEP-2002 (first entry)
XX
DE Human PSNA cDNA #8.
XX
KM Human; prostate specific protein; PSP; prostate specific nucleic acid;
KM vaccine; transgenic; prostate cancer; gene therapy; transgenic animal;
KM PSNA; ss.
XX
OS Homo sapiens.
XX
PN WO200224718-A1.
XX
PD 28-MAR-2002.
XX
PF 19-SEP-2001; 2001WO-US029386.
XX
PR 19-SEP-2000; 2000US-0233746P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Sun Y, Recipon H, Cafferty R, Ali S;
XX
DR WPI; 2002-471216/50.
XX
PT Novel isolated prostate specific polypeptide useful for identifying,
PT diagnosing, monitoring, staging, imaging, and treating prostate cancer
PT and non-cancerous disease states in prostate.
XX
PS Claim 1; Page 187; 210pp; English.
XX
CC The invention relates to prostate specific proteins (PSP) and prostate
CC specific nucleic acids (PSNA). Sequences of the invention are useful for
CC identifying, diagnosing, monitoring, staging, imaging and treating
CC prostate cancer and non-cancerous disease states in prostate. They are
CC also useful for producing engineered prostate tissue for treatment and
CC research. The PSNA sequences are used in gene therapy and for producing
CC transgenic animals and cells. The invention is also used as vaccines. The
CC present sequence is human PSNA cDNA
XX
SQ Sequence 1418 BP; 418 A; 335 C; 285 G; 380 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,46e-33 Length: 1418
Score: 307.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-031-158-14 (1-58) x AAD38828 (1-1418)

```

```

QY 1 MetGlnMetPheProPheSerProLeuPhePhePheLeuGlnLeuLeuYagInSerSer 20
   |||
DB 449 ATGCAGATGTTTCCCAAGCCACGATTTTCTTCTTCATGTCGAAACAAAGCTCC 508
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40
   |||
DB 509 AGAAGGCTGGACATACCTTTGCTCTTGAGAAATTTTCCCTGATGTATTAAATAC 568
QY 41 IleglyLysLysArgArgAlaThrArgPheTTPaPProArgArgGlyThrPro 58
   |||
DB 569 ATTGGCAAGAAAAGAGACCAACGATTCGGGATCCAGAGGGGAACACCA 622

RESULT 17
ADE08798/C
ID ADE08798 standard; DNA; 2658 BP.
XX
AC ADE08798;
XX
DT 29-JAN-2004 (first entry)
XX
DE Novel DNA-related contig nucleotide sequence #42.
XX
KM novel gene; novel protein; tissue marker; molecular weight marker;
KM chromosome marker; genetic disorder; contig; ds.
XX
OS Unidentified.
XX
PN WO2003054152-A2.
XX
PD 03-JUL-2003.
XX
PF 10-DEC-2002; 2002WO-US039555.
XX
PR 10-DEC-2001; 2001US-0339739P.
XX
PR 11-DEC-2001; 2001US-0339453P.
XX
PR 14-MAR-2002; 2002US-0365091P.
XX
PR 14-MAR-2002; 2002US-0365384P.
XX
PR 12-APR-2002; 2002US-0372381P.
XX
PR 12-APR-2002; 2002US-0372615P.
XX
PR 22-APR-2002; 2002US-0012855P.
XX
PR 24-APR-2002; 2002US-0376045P.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Asundi V, Goodrich RM, Ren F, Zhang J, Zhao QA, Wang J;
PI Ghosh M, Xue AJ, Wehrman T, Meng G, Zhou P, Drmanac RT, Wang Z;
PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;
XX
DR WPI; 2003-569235/53.
XX
PT New polynucleotides, useful for expressing recombinant proteins for
PT analysis, characterization or therapeutic use, or as markers for tissues
PT in which the corresponding protein is preferentially expressed.
XX
PS Disclosure; SEQ ID NO 1864; 1177pp; English.
XX
CC The invention comprises the amino acid and coding sequences of novel
CC proteins. The DNA and protein sequences of the invention are useful as:
CC markers for tissues in which the corresponding protein is preferentially
CC expressed; as molecular weight markers on gels; as chromosome markers or
CC tags; to identify chromosomes or to map related gene positions; and to
CC compare with endogenous DNA sequences in patients to identify potential
CC genetic disorders. The present DNA sequence was used in the
CC exemplification of the invention.
XX
SQ Sequence 2658 BP; 745 A; 585 C; 578 G; 750 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5,74e-33 Length: 2658
Score: 307.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0

```

DB: 10 Gaps: 0

US-10-031-158-14 (1-58) x ADE08798 (1-2658)

QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20

DB 2381 ATGCAGATGTTTCCCCCAAGCCACTATTTTCTTCAATTCGAAACAAAGCTCC 2322

QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAnpPheSerLeuMetLeuLeuArgTyr 40

DB 2331 AGAAGGCTGGAACATCACTTTGTCTTCTTGAGAAATTTTCCCTAATGTTATTAAGATAC 2262

QY 41 lIleGlyLysLysArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58

DB 2261 ATTGGCAAGAAAGAAAGACACAGATTCGGATCCAGAGGGGAACACCA 2208

RESULT 18

ABV25013/C

ID ABV25013 standard; cDNA; 466 BP.

XX

AC ABV25013;

XX

DT 16-SEP-2002 (first entry)

XX

DE Human prostate expression marker CDNA 25004.

XX

KM Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX

KM pharmacogenomic marker; gene; ss.

XX

OS Homo sapiens.

XX

PN WO200160860-A2.

XX

PD 23-AUG-2001;

XX

PF 20-FEB-2001; 2001WO-US005171.

XX

PR 17-FEB-2000; 2000US-0183319P.

XX

PR 16-MAR-2000; 2000US-0189862P.

XX

PR 25-MAY-2000; 2000US-0207454P.

XX

PR 09-JUN-2000; 2000US-0211314P.

XX

PR 18-JUL-2000; 2000US-0219007P.

XX

PR 13-DEC-2000; 2000US-0255281P.

XX

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Endege WO, Monahan JE;

XX

DR WPI; 2001-662795/76.

XX

PT Novel isolated nucleic acid molecule associated with cancerous state of

PT prostate cells and correlating with presence of prostate cancer, useful

PT for detecting presence of prostate cancer, stage of prostate cancer.

XX

PS Claim 1; Page 4846; 11750pp; English.

XX

CC The invention relates to an isolated nucleic acid molecule (1) comprising

CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

CC specification or its complement. (1) is useful for: (a) assessing whether

CC a patient is afflicted with prostate cancer; (b) monitoring the

CC progression of prostate cancer in a patient; (c) assessing the efficacy

CC of a test compound to inhibit prostate cancer in a patient; (d) assessing

CC the efficacy of a therapy for inhibiting prostate cancer in a patient;

CC (e) selecting a composition for inhibiting prostate cancer in a patient;

CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)

CC determining whether prostate cancer has metastasized in a patient; (h)

CC assessing the aggressiveness or indolence of prostate cancer in a patient

CC ; (1) is also useful as a pharmacodynamic or pharmacogenomic marker

XX

SQ Sequence 486 BP; 117 A; 94 C; 101 G; 174 T; 0 U; 0 Other;

Alignment Scores: 1.55e-33 Length: 486

Pred. No.:

Score: 304.00 Matches: 57

Percent Similarity: 100.00% Conservatve: 1

Best Local Similarity: 98.28% Mismatches: 0

Query Match: 99.02% Indels: 0

DB: 5 Gaps: 0

US-10-031-158-14 (1-58) x ABV25013 (1-486)

QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20

DB 411 ATGCAGATGTTTCCCCCAAGCCACTATTTTCTTCAATTCGAAACAAAGCTCC 352

QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAnpPheSerLeuMetLeuLeuArgTyr 40

DB 351 AGAAGGCTGGAACATCACTTTGTCTTCTTGAGAAATTTTCCCTAATGTTATTAAGATAC 292

QY 41 lIleGlyLysLysArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58

DB 291 ATTGGCAAGAAAGAAAGACACAGATTCGGATCCAGAGGGGAACACCA 238

RESULT 19

ABA47985

ID ABA47985 standard; DNA; 316 BP.

XX

AC ABA47985;

XX

DT 01-FEB-2002 (first entry)

XX

DE Human breast cell single exon nucleic acid probe #6680.

XX

KM Human; microarray; single exon probe; gene expression; breast; disease;

XX

KM cancer; ss.

XX

OS Homo sapiens.

XX

PN WO200157271-A2.

XX

PD 09-AUG-2001.

XX

PF 30-JAN-2001; 2001WO-US000662.

XX

PR 04-FEB-2000; 2000US-0180312P.

XX

PR 26-MAY-2000; 2000US-0207456P.

XX

PR 30-JUN-2000; 2000US-00608408.

XX

PR 03-AUG-2000; 2000US-0063236P.

XX

PR 21-SEP-2000; 2000US-0234687P.

XX

PR 27-SEP-2000; 2000US-0236359P.

XX

PR 04-OCT-2000; 2000GB-00024263.

XX

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX

DR WPI; 2001-496933/54.

XX

PT New spatially-addressable set of single exon nucleic acid probes, useful

PT for measuring gene expression in sample derived from human breast,

PT comprises number of single exon nucleic acid probes.

XX

PS Claim 4; SEQ ID NO 6680; 327pp + Sequence Listing; English.

XX

CC The invention relates to a spatially-addressable set of single exon

CC nucleic acid probes for measuring gene expression in a sample derived

CC from human breast and BR 474 cells. The method involves contacting the

CC probes with a collection of detectably labelled nucleic acids derived

CC from mRNA of human breast, and then measuring the label bound to each

CC probe of the microarray. The probes are useful for verifying the

CC expression of regions of genomic DNA predicted to encode proteins. They

CC are useful for gene discovery, and for determining predisposition and/or

CC prognosing breast disease. Gene expression analysis is useful for

CC assessing the toxicity of chemical agents on cells. The microarray of

CC this invention presents a far greater diversity of probes for measuring

CC gene expression, with far less bias than expressed sequence tag

CC microarrays. The method is suitable for rapid production of functional  
CC information from genomic sequence. The present sequence is a single exon  
CC nucleic acid probe of the invention. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pcc\_sequences

XX SQ Sequence 316 BP; 112 A; 65 C; 62 G; 77 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.67e-33	Length:	316
Score:	302.00	Matches:	57
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	98.37%	Indels:	0
DB:	4	Gaps:	0

US-10-031-158-14 (1-58) x ABA47985 (1-316)

QY 2 GlnMetPheProSerProLeuPhePheLeuGlnLeuLysGlnSerSerArg 21  
|||  
DB 2 CAGATGTTTCCCCCAAGCCCACTATTTCTTCTTCAATGCTGAACAAAGCTCCACA 61  
|||  
QY 22 ArgLeuGlnHisThrPheValPheLeuArgaenPheSerLeuMetLeuLysArgTyrIle 41  
|||  
DB 62 AGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAGATACATT 121  
|||  
QY 42 GtLYsLYsArgArgAlaThrArgPheTTPaerProArgArgGlyThrPro 58  
|||  
DB 122 GGCNAAGAAAGAGAGACACGATTCGGATCCAGAGGGGAACCA 172  
|||

RESULT 20

AAK40015  
ID AAK40015 standard; DNA; 316 BP.

XX AAK40015;

XX DT 06-NOV-2001 (first entry)

XX DE Human bone marrow expressed single exon probe SEQ ID NO: 14572.

XX KW Human; bone marrow expressed exon; gene expression analysis; probe;

XX KM microarray; cancer; leukemia; lymphoma; myeloma; ss.

XX OS Homo sapiens.

XX PN MO200157276-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001MO-US000668.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488900/53.

XX PT Human genome-derived single exon nucleic acid probes useful for analyzing  
XX gene expression in human bone marrow.  
XX  
XX Example 4; SEQ ID NO 14572; 658bp + Sequence Listing; English.  
XX  
XX The present invention provides a number of single exon nucleic acid  
XX probes which are derived from genomic sequences expressed in the human  
CC

CC bone marrow. They can be used to measure gene expression in bone marrow  
CC samples, which may enable the improved diagnosis and treatment of cancers  
CC such as lymphoma, leukemia and myeloma. The present sequence is one of  
CC the probes of the invention

XX SQ Sequence 316 BP; 112 A; 65 C; 62 G; 77 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.67e-33	Length:	316
Score:	302.00	Matches:	57
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	98.37%	Indels:	0
DB:	4	Gaps:	0

US-10-031-158-14 (1-58) x AAK40015 (1-316)

QY 2 GlnMetPheProSerProLeuPhePheLeuGlnLeuLysGlnSerSerArg 21  
|||  
DB 2 CAGATGTTTCCCCCAAGCCCACTATTTCTTCTTCAATGCTGAACAAAGCTCCACA 61  
|||  
QY 22 ArgLeuGlnHisThrPheValPheLeuArgaenPheSerLeuMetLeuLysArgTyrIle 41  
|||  
DB 62 AGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAGATACATT 121  
|||  
QY 42 GtLYsLYsArgArgAlaThrArgPheTTPaerProArgArgGlyThrPro 58  
|||  
DB 122 GGCNAAGAAAGAGAGACACGATTCGGATCCAGAGGGGAACCA 172  
|||

RESULT 21

ABS14101  
ID ABS14101 standard; DNA; 316 BP.

XX ABS14101;

XX DT 19-AUG-2002 (first entry)

XX DE Human genome-derived single exon probe ORF from lung SEQ ID NO 14092.

XX KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;

XX KM chronic obstructive pulmonary disease; interstitial lung disease;

XX KW familial idiopathic pulmonary fibrosis; neurofibromatosis;

XX KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;

XX KW Hereditary spherocytosis; sarcoidosis; pulmonary haemosiderosis;

XX KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;

XX KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;

XX KW primary ciliary dyskinesia; pulmonary hypertension;

XX KW hyaline membrane disease; open reading frame; ORF.

XX OS Homo sapiens.

XX PN MO200186003-A2.

XX PD 15-NOV-2001.

XX PF 30-JAN-2001; 2001MO-US000665.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2002-114183/15.

XX PT Spatially-addressable set of single exon nucleic acid probes, used to  
XX measure gene expression in human lung samples.  
XX  
XX



XX ABL65007;  
 AC 15-MAY-2002 (first entry)  
 XX  
 DT Lung cancer related gene sequence SEQ ID NO:3344.  
 XX  
 DE Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
 KW cytostatic; gene therapy; antineoplastic; Wilms tumour; adenocarcinoma;  
 KM gene; ds.  
 XX Homo sapiens.  
 OS  
 XX MO200194629-A2.  
 PN  
 XX 13-DEC-2001.  
 PD  
 XX  
 PF 30-MAY-2001; 2001WO-US010838.  
 XX  
 PR 05-JUN-2000; 2000US-0209473P.  
 PR 05-JUN-2000; 2000US-0209531P.  
 PR 18-SEP-2000; 2000US-0233133P.  
 PR 18-SEP-2000; 2000US-0233617P.  
 PR 20-SEP-2000; 2000US-0234009P.  
 PR 20-SEP-2000; 2000US-0234034P.  
 PR 20-SEP-2000; 2000US-0234052P.  
 PR 22-SEP-2000; 2000US-0234509P.  
 PR 22-SEP-2000; 2000US-0234567P.  
 PR 25-SEP-2000; 2000US-0234923P.  
 PR 25-SEP-2000; 2000US-0234924P.  
 PR 25-SEP-2000; 2000US-0235077P.  
 PR 25-SEP-2000; 2000US-0235082P.  
 PR 25-SEP-2000; 2000US-0235134P.  
 PR 25-SEP-2000; 2000US-0235280P.  
 PR 26-SEP-2000; 2000US-0235637P.  
 PR 26-SEP-2000; 2000US-0235638P.  
 PR 27-SEP-2000; 2000US-0235711P.  
 PR 27-SEP-2000; 2000US-0235720P.  
 PR 27-SEP-2000; 2000US-0235840P.  
 PR 27-SEP-2000; 2000US-0235863P.  
 PR 28-SEP-2000; 2000US-0236028P.  
 PR 28-SEP-2000; 2000US-0236032P.  
 PR 28-SEP-2000; 2000US-0236033P.  
 PR 28-SEP-2000; 2000US-0236034P.  
 PR 28-SEP-2000; 2000US-0236109P.  
 PR 28-SEP-2000; 2000US-0236111P.  
 PR 29-SEP-2000; 2000US-0236642P.  
 PR 29-SEP-2000; 2000US-0236691P.  
 PR 02-OCT-2000; 2000US-0237172P.  
 PR 02-OCT-2000; 2000US-0237173P.  
 PR 02-OCT-2000; 2000US-0237278P.  
 PR 02-OCT-2000; 2000US-0237294P.  
 PR 02-OCT-2000; 2000US-0237295P.  
 PR 02-OCT-2000; 2000US-0237316P.  
 PR 03-OCT-2000; 2000US-0237425P.  
 PR 03-OCT-2000; 2000US-0237598P.  
 PR 03-OCT-2000; 2000US-0237604P.  
 PR 03-OCT-2000; 2000US-0237606P.  
 PR 03-OCT-2000; 2000US-0237608P.  
 PR 01-NOV-2000; 2000US-0244867P.  
 PR 01-NOV-2000; 2000US-0245084P.  
 XX  
 PA (AVALON PHARM.  
 XX  
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
 PI Sopet DR, Weaver Z;  
 XX  
 DR MPI, 2002-188264/24.  
 XX  
 PT Screening for anti-neoplastic agent involves exposing cells to a chemical  
 PT agent to be tested for anti-neoplastic activity, and determining a change  
 PT in expression of a gene of a signature gene set.

XX Claim 1; SEQ ID NO 3344; 44bp; English.  
 XX  
 CC The present invention describes a method (M1) for screening for an anti-  
 CC neoplastic agent. The method involves exposing cells to a chemical agent  
 CC to be tested for anti-neoplastic activity, determining a change in (I)  
 CC expression of at least one gene (I) of a signature gene set, where (I)  
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664  
 CC to ABL70110), or is at least 95% identical to (S), where a change in  
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
 CC activity and can be used in gene therapy. M1 can be used for screening an  
 CC anti-neoplastic agent, and can be used for producing a product which is  
 CC the data collected with respect to the anti-neoplastic agent as a result  
 CC of M1, and the data is sufficient to convey the chemical structure and/or  
 CC properties of the agent. M1 can be used in the treatment of cancer such  
 CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,  
 CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell  
 CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous  
 CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms  
 CC tumour  
 XX  
 XX Sequence 1421 BP; 419 A; 325 C; 292 G; 385 T; 0 U; 0 Other;  
 SO  
 Alignment Scores:  
 Pred. No.: 1.74e-31 Length: 1421  
 Score: 294.00 Matches: 55  
 Percent Similarity: 98.28% Conservative: 2  
 Best Local Similarity: 94.83% Mismatches: 1  
 Query Match: 95.77% Indels: 0  
 DB: Gaps: 0  
 US-10-031-158-14 (1-58) x ABL65007 (1-1421)  
 QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuysGlnSerSer 20  
 DB 423 ATGCAAGATGTTTCCGCCAAGCCACTATTTCTTCCTTCATTTGCTGAACAAACCTC 482  
 QY 21 ATGATGLeuGlnIleThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40  
 DB 483 ACAAGCGCTGAAACATACCTTGTCTTCTGAGAAATTTTCCACATATATTAAAGATAC 542  
 QY 41 TTEGTYLVLYVLYARGATGATGATAThrArgPheTTPAARProArgArgGlyThrPro 58  
 DB 543 ATTGGCAAGAAAGAAAGACACGATTTCTGGATCCAGAGAGGAGAACCA 596  
 RESULT 24  
 ADE06987  
 ID ADE06987 standard; DNA; 1560 BP.  
 AC ADE06987;  
 AC  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Novel coding sequence (useful for identifying genetic disorders) #53.  
 XX  
 KW novel gene; novel protein; tissue marker; molecular weight marker;  
 KW chromosome marker; genetic disorder; gene; ds.  
 OS Unidentified.  
 XX  
 PN WO2003054152-A2.  
 PD  
 PD 03-JUL-2003.  
 XX  
 PF 10-DEC-2002; 2002WO-US039555.  
 XX  
 XX 10-DEC-2001; 2001US-0339739P.  
 PR 11-DEC-2001; 2001US-0339453P.  
 PR 14-MAR-2002; 2002US-0365091P.  
 PR 14-MAR-2002; 2002US-0365384P.  
 PR 12-APR-2002; 2002US-0372381P.  
 PR 12-APR-2002; 2002US-0372615P.  
 PR 22-APR-2002; 2002US-00128558.

PR 24-APR-2002; 2002US-0376045P.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J,  
PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Dimanac RT, Wang Z,  
PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ,  
XX  
XX WPI: 2003-569235/53.  
DR P-PSDB; ADB07898.  
XX  
XX New polynucleotides, useful for expressing recombinant proteins for  
PT analysis, characterization or therapeutic use, or as markers for tissues  
PT in which the corresponding protein is preferentially expressed.  
XX  
XX Claim 1; SEQ ID NO 53; 1177bp; English.  
XX  
XX The invention comprises the amino acid and coding sequences of novel  
CC proteins. The DNA and protein sequences of the invention are useful as:  
CC markers for tissues in which the corresponding protein is preferentially  
CC expressed; as molecular weight markers on gels; as chromosome markers or  
CC tags; to identify chromosomes or to map related gene positions; and to  
CC compare with endogenous DNA sequences in patients to identify potential  
CC genetic disorders. The present DNA sequence represents a gene of the  
CC invention.  
XX  
XX Sequence 1560 BP; 444 A; 363 C; 322 G; 431 T; 0 U; 0 Other;  
SQ  
Alignment Scores:  
Pred. No.: 1.97e-31 Length: 1560  
Score: 294.00 Matches: 55  
Percent Similarity: 98.28% Conservative: 2  
Best Local Similarity: 95.83% Mismatches: 1  
Query Match: 95.77% Indels: 0  
DB: 10 Gaps: 0  
US-10-031-158-14 (1-58) x ADB06987 (1-1560)  
QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20  
DB 559 ATGCAGATGTTTCCCAAGCCCACTATTCTTCTTCAATGCTGAACAAACCTCC 618  
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAnpSerLeuMetLeuLeuArgTyr 40  
DB 619 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCAATATTATTAAGATAC 678  
QY 41 IleGlyLysLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58  
DB 679 ATTGGCAAGAAAGAGCAACAGATTCTGTGATCCAGAGGGAACACCA 732  
RESULT 25  
ABZ35381  
ID ABZ35381 standard; cDNA; 1799 BP.  
XX  
XX AC ABZ35381;  
XX  
XX DT 05-FEB-2003 (first entry)  
XX  
XX Human gene expression profile polynucleotide SEQ ID NO 492.  
XX  
XX Human; artery; endothelium; umbilical; vein; aorta; pulmonary artery;  
KW bronchial epithelium; prostate; muscle; lung fibroblast; osteoblast;  
KW tumour; macrophage; genome mapping; antibiotic; antiviral; anti-fungal;  
KW gene expression; gene; ss.  
XX  
XX Homo sapiens.  
XX OS  
XX WO200274979-A2.  
XX  
XX PD 26-SEP-2002.  
XX  
XX PF 20-MAR-2002; 2002WO-US008456.  
XX

PR 20-MAR-2001; 2001US-0276947P.  
XX  
XX (ORTH) ORTHO CLINICAL DIAGNOSTICS INC.  
XX  
XX Wan J, Wang Y;  
PI  
XX  
XX WPI: 2002-740862/80.  
DR  
XX  
XX New gene expression profile generated from primary, endothelial,  
PT epithelial, and muscle cell types, useful for identifying disease  
PT pathologies involving alterations of gene expression, e.g. cancer.  
XX  
XX Example 3; Page 634-635; 850bp; English.  
XX  
XX The invention relates to a gene expression profile comprising one or more  
CC genes (ABZ34889-ABZ35692) and generated from a cell type. The cell type  
CC is a coronary artery endothelium, umbilical artery or vein endothelium,  
CC aortic endothelium, dermal microvascular endothelium, pulmonary artery  
CC endothelium, myometrium microvascular endothelium, keratinocyte  
CC epithelium, bronchial epithelium, mammary epithelium, prostate  
CC epithelium, renal cortical epithelium, renal proximal tubule epithelium,  
CC small airway epithelium, renal epithelium, umbilical artery smooth  
CC muscle, neonatal dermal fibroblast, pulmonary artery smooth muscle,  
CC dermal fibroblast, neural progenitor cells, skeletal muscle, astrocytes,  
CC aortic smooth muscle, mesangial cells, coronary artery smooth muscle,  
CC bronchial smooth muscle, uterine smooth muscle, lung fibroblast,  
CC osteoblasts or prostate stromal cell. The gene expression profile is used  
CC for determining the level of RNA expression for a sample, determining the  
CC phenotype of a cell and distinguishing cell types. The gene or a protein  
CC expression profile is useful in identifying disease pathologies involving  
CC alterations of gene expression. The assessment of expression profiles may  
CC provide meaningful information with respect to tumour type and stage,  
CC treatment methods, and prognosis. The gene or protein expression profile  
CC may also be used for creating microarrays. The microarray is useful for  
CC genetic and physical mapping of genomes, DNA sequencing, genetic or  
CC medical diagnosis, genotyping of organisms, confirming cell or tissue  
CC identifications and in identifying promising antibiotics, antiviral or  
CC antifungal agents  
SQ  
Sequence 1799 BP; 499 A; 434 C; 396 G; 470 T; 0 U; 0 Other;  
XX  
Alignment Scores:  
Pred. No.: 2.39e-31 Length: 1799  
Score: 294.00 Matches: 58  
Percent Similarity: 98.31% Conservative: 0  
Best Local Similarity: 98.31% Mismatches: 0  
Query Match: 95.77% Indels: 1  
DB: 6 Gaps: 0  
US-10-031-158-14 (1-58) x ABZ35381 (1-1799)  
QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20  
DB 845 ATGCAGATGTTTCCCAAGCCCACTATTCTTCTTCAATGCTGAACAAAGCTCC 904  
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAnpSerLeuMetLeuLeuArgTyr 40  
DB 905 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCAATATTATTAAGATAC 964  
QY 41 IleGlyLysLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58  
DB 965 ATTGGCAAGAAAGAGCAACAGATTCTGTGATCCAGAGGGAACACCA 1019  
RESULT 26  
AAQ37617  
ID AAQ37617 standard; cDNA; 1080 BP.  
XX  
XX AC AAQ37617;  
XX  
XX DT 25-MAR-2003 (revised)  
XX  
XX DT 21-JUN-1993 (first entry)  
XX  
XX TCR gamma protein (cell line MOLT-13).  
DE

```

XX T cell antigen receptor; TCR; IDP2; PBER; MOLT-3; immunoglobulin;
KM cell surface; delta chain; Form 2bc; ss.
XX Synthetic.
XX Key Location/Qualifiers
FH CDS 37..1008
FT /tag= a
FT /label= gamma TCR
FT /note= "comprises S, V, N, J, C1, C1B, C1C and C1I
FT regions"
XX US185250-A.
XX 09-FEB-1993.
XX 13-JAN-1989; 89US-00297661.
XX 03-JUL-1986; 86US-00882100.
XX 19-FEB-1987; 87US-00016252.
XX 29-OCT-1987; 87US-00115256.
XX 29-APR-1988; 88US-00187698.
XX (TCEL-) T CEL SCI INC.
XX (DAND ) DANA FARBER CANCER INST INC.
XX Brenner MB, Seidman J, Strominger JL, Ip SH, Krangel MS, Band H,
XX WPI; 1993-067162/08.
XX F-PSDB; AAR32906.
XX Gamma, delta T-cell antigen receptor heterodimer - used to diagnose
XX immune deficiencies e.g. AIDS and stimulate lymphocytes.
XX Disclosure; Fig 10A-B; 99pp; English.
XX Example 11 describes the three forms of the human TCR gamma/delta. The
XX new form of the human gamma/delta TCR consists of a 40 kD TCR gamma
XX glycoprotein (termed Form 2bc) noncovalently associated with a TCR delta
XX chain. Clone M13k corresponds to a full length, in frame, gamma TCR
XX transcript. (Updated on 25-MAR-2003 to correct PF field.)
XX SQ Sequence 1080 BP; 324 A; 260 C; 231 G; 265 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 4.45e-31 Length: 1080
XX Score: 290.00 Matches: 54
XX Percent Similarity: 98.28% Conservative: 3
XX Best Local Similarity: 93.10% Mismatches: 1
XX Query Match: 94.46% Indels: 0
XX DB: 2 Gaps: 0
XX
XX US-10-031-158-14 (1-58) x AAQ37617 (1-1080)
XX
XX QY 1 MetGlnMePheProSerProLeuPhePheLeuGlnLeuLeuYsGlnSerSer 20
XX |||ATGCAAGATGTTTCCCAAGCCACATATTTCTTCCTTGATGCTGGAACAAACTCC 511
XX
XX DB 452 ATGCAAGATGTTTCCCAAGCCACATATTTCTTCCTTGATGCTGGAACAAACTCC 511
XX
XX QY 21 ArgArgLeuGlnIuHsrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTYr 40
XX |||AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCAGATATTATTAAGATAC 571
XX
XX DB 512 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCAGATATTATTAAGATAC 571
XX
XX QY 41 ILGLYLYELYSARGALATHARGPHEITPASPProARGIYThrPro 58
XX |||ATTGGCAAGAAAGAAAGACACACGATTCGGATCCGAGAGGGGAAACACCA 625
XX
XX DB 572 ATTGGCAAGAAAGAAAGACACACGATTCGGATCCGAGAGGGGAAACACCA 625
XX
XX RESULT 27
XX ADE08798 standard; DNA; 2658 BP.
XX AC ADE08798;
XX

```

```

DT 29-JAN-2004 (first entry)
XX
XX DE Novel DNA-related contig nucleotide sequence #42.
XX
XX KW novel gene; novel protein; tissue marker; molecular weight marker;
XX chromosome marker; genetic disorder; contig; ds.
XX
XX OS Unidentified.
XX
XX PN WO2003054152-A2.
XX
XX 03-JUL-2003.
XX
XX 10-DEC-2002; 2002WO-US039555.
XX
XX 11-DEC-2001; 2001US-0339739P.
XX 11-DEC-2001; 2001US-0339453P.
XX 14-MAR-2002; 2002US-0365091P.
XX 14-MAR-2002; 2002US-0365384P.
XX 12-APR-2002; 2002US-0372381P.
XX 12-APR-2002; 2002US-0372615P.
XX 22-APR-2002; 2002US-00128558.
XX 24-APR-2002; 2002US-0376045P.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
XX Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Dymnac RT, Wang Z;
XX Ma Y, Wang D, Chen R, Xu C, Boyle BJ;
XX WPI; 2003-569235/53.
XX
XX New polynucleotides, useful for expressing recombinant proteins for
XX PT analysis, characterization or therapeutic use, or as markers for tissues
XX PT in which the corresponding protein is preferentially expressed.
XX
XX PS Disclosure; SEQ ID NO 1864; 1177pp; English.
XX
XX CC The invention comprises the amino acid and coding sequences of novel
XX CC proteins. The DNA and protein sequences of the invention are useful as:
XX CC markers for tissues in which the corresponding protein is preferentially
XX CC expressed; as molecular weight markers on gels; as chromosome markers or
XX CC tags; to identify chromosomes or to map related gene positions; and to
XX CC compare with endogenous DNA sequences in patients to identify potential
XX CC genetic disorders. The present DNA sequence was used in the
XX CC exemplification of the invention.
XX
XX SQ Sequence 2658 BP; 745 A; 585 C; 578 G; 750 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 1.5e-30 Length: 2658
XX Score: 290.00 Matches: 54
XX Percent Similarity: 98.28% Conservative: 3
XX Best Local Similarity: 93.10% Mismatches: 1
XX Query Match: 94.46% Indels: 0
XX DB: 10 Gaps: 0
XX
XX US-10-031-158-14 (1-58) x ADE08798 (1-2658)
XX
XX QY 1 MetGlnMePheProSerProLeuPhePheLeuGlnLeuLeuYsGlnSerSer 20
XX |||ATGCAAGATGTTTCCCAAGCCACATATTTCTTCCTTGATGCTGGAACAAACTCC 556
XX
XX DB 497 ATGCAAGATGTTTCCCAAGCCACATATTTCTTCCTTGATGCTGGAACAAACTCC 556
XX
XX QY 21 ArgArgLeuGlnIuHsrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTYr 40
XX |||AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCAGATATTATTAAGATAC 616
XX
XX DB 557 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCAGATATTATTAAGATAC 616
XX
XX QY 41 ILGLYLYELYSARGALATHARGPHEITPASPProARGIYThrPro 58
XX |||ATTGGCAAGAAAGAAAGACACACGATTCGGATCCGAGAGGGGAAACACCA 670
XX
XX DB 617 ATTGGCAAGAAAGAAAGACACACGATTCGGATCCGAGAGGGGAAACACCA 670
XX
XX RESULT 28
XX ADP10449
XX

```



```

QY      1 MetGlnMetPheProSerProLeuPhePhePheLeuGlnLeuLeuYsgInserSer 20
DB      553 ATGCAGATGTTTCCCCCAAGCCCACTATTTCCTTCCTTCGATTCGTGAAACAAACTCC 612
QY      21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40
DB      613 AGAAGGCTGGAACATATCTTGTCTTCTTGAGAAATTTTCCAGATATTATTAAAGATAC 672
QY      41 IleglyLVsLYsArgArgAlaThrArgPheTTPAspProArgArgGlyThrPro 58
DB      673 ATTGGCAAGAAAAGAGAGCAACACGATTCGTGGATCCAGAGGGGAAACACCA 726
RESULT 30
AAN91698
ID      AAN91698 standard; DNA; 1080 BP.
XX
XX      AAN91698;
AC
XX
XX      25-MAR-2003 (revised)
DT      31-OCT-2002 (revised)
DT      26-APR-1990 (first entry)
XX
XX      Sequence of MOLT-13 cells gamma T cell antigen receptor (Form 2bc).
DE
XX      Human gamma T cell antigen receptor; gamma TCR; MOLT-13 cells;
XX      clone M13k.
KM
XX
XX      Homo sapiens.
OS
XX
XX      Key
FH      Location/Qualifiers
FT      37..1008
FT      misc_signal
FT      37..46
FT      misc_RNA
FT      80..84
FT      misc_RNA
FT      391..394
FT      misc_RNA
FT      398..406
FT      misc_RNA
FT      439..448
FT      misc_RNA
FT      770..778
FT      misc_RNA
FT      817..826
FT      misc_RNA
FT      871..880
FT      misc_RNA
FT      /product= "CIII constant region"
XX
XX      WO890396-A.
XX
XX      05-MAY-1989.
XX
XX      28-OCT-1988; 88WO-US003869.
XX
XX      29-OCT-1987; 87US-00115256.
XX      29-APR-1988; 88US-00187698.
XX
XX      (TCELL-) TCELL SCIENCES INC.
XX      (DAND-) DANA-FABBER CANCER INST.
XX      (HARD) HARVARD COLLEGE.
XX
XX      Brenner MB, Ip SH, Seidman J;
XX      WPI; 1989-150856/20.
DR

```

```

DR      P-PSDB; AAP91885.
XX
XX      Human gamma, delta T cell antigen receptor - studied by isolation of
PT      polypeptide(s), nucleic acids and monoclonal antibodies reactive with
PT      epitope(s).
XX
XX      Fig 14; Page 7; 104pp; English.
XX
XX      The CDS (tag b) codes for human gamma T cell antigen receptor polypeptide
CC      termed Form 2bc. The sequence can be DNA or RNA. It was prepd. as
CC      follows: a cDNA library was prepd. from MOLT-13 poly(A)+ RNA in lambda
CC      g10 vector. This library was amplified on the E.coli strain C600 Hfl and
CC      screened. cDNA clone M13k was selected for sequencing. It corresponds to
CC      a full length, in-frame gamma T cell antigen receptor transcript.
CC      (Updated on 31-OCT-2002 to add missing 06 field.) (Updated on 25-MAR-2003
CC      to correct PR field.) (Updated on 25-MAR-2003 to correct PA field.)
XX
XX      Sequence 1080 BP; 324 A; 257 C; 233 G; 266 T; 0 U; 0 Other;
SQ

```

Alignment Scores:

Pred. No.:	1,63e-29	Length:	1080
Score:	279.00	Matches:	53
Percent Similarity:	96.55%	Conservative:	3
Best Local Similarity:	91.38%	Mismatches:	2
Query Match:	90.88%	Indels:	0
DB:	1	Gaps:	0

```

US-10-031-158-14 (1-58) x AAN91698 (1-1080)
QY      1 MetGlnMetPheProSerProLeuPhePhePheLeuGlnLeuLeuYsgInserSer 20
DB      452 ATGCAGATGTTTCCCCCAAGCCCACTATTTCCTTCCTTCGATTCGTGAAACAAACTCC 511
QY      21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40
DB      512 AGAAGGCTGGAACATATCTTGTCTTCTTGAGAAATTTTCCAGATATTATTAAAGATAG 571
QY      41 IleglyLVsLYsArgArgAlaThrArgPheTTPAspProArgArgGlyThrPro 58
DB      572 ATTGGCAAGAAAAGAGAGCAACACGATTCGTGGATCCAGAGGGGAAACACCA 625
RESULT 31
ABV37856
ID      ABV37856 standard; cDNA; 533 BP.
XX
XX      ABV37856;
AC
XX
XX      16-SEP-2002 (first entry)
DT
XX
XX      Human prostate expression marker cDNA 37847.
DE
XX      Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KM      pharmacogenomic marker; gene; ss.
XX
XX      Homo sapiens.
OS
XX
XX      WO200160860-A2.
XX
XX      23-AUG-2001.
XX
XX      20-FEB-2001; 2001WO-US005171.
XX
XX      17-FEB-2000; 2000US-0183319P.
XX      16-MAR-2000; 2000US-0189862P.
XX      25-MAY-2000; 2000US-0207454P.
XX      09-JUN-2000; 2000US-0211314P.
XX      18-JUL-2000; 2000US-0219007P.
XX      13-DEC-2000; 2000US-025281P.
XX
XX      (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX      Schlegel R, Endege WO, Monahan JE;
XX

```

DR WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1; Page 7745; 11750bp; English.

XX The invention relates to an isolated nucleic acid molecule (1) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (1) is useful for: (a) assessing whether  
CC a patient is afflicted with prostate cancer; (b) monitoring the  
CC progression of prostate cancer in a patient; (c) assessing the efficacy  
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
CC determining whether prostate cancer has metastasized in a patient; (h)  
CC assessing the aggressiveness or indolence of prostate cancer in a patient  
CC ; (1) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX

SQ Sequence 533 BP; 180 A; 116 C; 103 G; 132 T; 0 U; 2 Other;

## Alignment Scores:

Pred. No.:	4.38e-27	Length:	533
Score:	259.00	Matches:	54
Percent Similarity:	94.74%	Conservative:	0
Best Local Similarity:	94.74%	Mismatches:	3
Query Match:	84.36%	Indels:	1
DB:	5	Gaps:	0

US-10-031-158-14 (1-58) x ABV37856 (1-533)

QY 2 GlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSerArg 21

DB 123 CAGATGTTTCCTC--AGCCCACTATTTCCTTCCTCAATGCTGAAACAAAGCTCCAGA 180

QY 22 ArgLeuGlnHisThrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 41

DB 181 AGCGTGGACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAAGATPAC 240

QY 42 GlyLysLysArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58

DB 241 GGCAAGAAAGAAAGACACACGATTCGTGATCCAGAGGGAACACCA 291

## RESULT 32

AAC04122

ID AAC04122 standard; cDNA; 460 BP.

XX AAC04122;

DT 06-OCT-2000 (first entry)

DE Human secreted protein 5' EST, SEQ ID NO: 8197.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

KW gene therapy; chromosome mapping; ss.

OS Homo sapiens.

PN EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-00200610.

XX 26-FEB-1999; 99US-0122487P.

XX (GEST ) GENSET.

PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for  
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.

XX Claim 1; SEQ ID NO 8197; 71bp + Sequence Listing; English.

XX The present sequence is one of a large number of 5' ESTs derived from  
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively  
CC identified within the present sequence. The 5' ESTs were prepared from  
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST  
CC sequences usually correspond mainly to the 3' untranslated region (UTR)  
CC of the mRNA because they are often obtained from oligo-dT primed cDNA  
CC libraries. Such ESTs are not well suited for isolating cDNA sequences  
CC derived from the 5' ends of mRNAs and even in those cases where longer  
CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'  
CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used  
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in  
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.  
CC They are used to obtain upstream regulatory sequences and to design  
CC expression and secretion vectors

SQ Sequence 460 BP; 162 A; 83 C; 90 G; 121 T; 0 U; 4 Other;

Alignment Scores:			
Pred. No.:	4.93e-26	Length:	460
Score:	251.00	Matches:	56
Percent Similarity:	96.55%	Conservative:	0
Best Local Similarity:	96.55%	Mismatches:	2
Query Match:	81.76%	Indels:	2
DB:	3	Gaps:	0

US-10-031-158-14 (1-58) x AAC04122 (1-460)

QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20

DB 200 ATGCAGATGTTTCCCAAGCCCACTATTTCCTTCCTCAATGCTGAAACAAAGCTC- 258

QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40

DB 259 AGAAGGCTGGAAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAAGATPAC 318

QY 41 IleGlyLysLysArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58

DB 319 ATTGGGAARARAAAGAGACACACGATC-TGGATCCAGAGGGAACACCA 371

## RESULT 33

AAF15623

ID AAF15623 standard; cDNA; 1155 BP.

XX AAF15623;

DT 13-MAR-2001 (first entry)

DE Human prostate cancer antigen nucleotide sequence SEQ ID NO:58.

XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;

KW neuroprotective; cytoskeletal; cardiovascular; immunomodulatory; muscular;

KW vulnereary; gastrointestinal; nephrotropic; antiinfective; gynaecological;

KW antibacterial; gene therapy; neural; immune; reproductive; renal;

KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;

KW wound; infectious disease; ss.

XX Homo sapiens.

XX WO200055174-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US005988.

XX 12-MAR-1999; 99US-0124270P.



XX 23-AUG-2001.  
PD  
XX  
PF 20-FEB-2001; 2001WO-US005171.  
XX  
PR 17-FEB-2000; 2000US-0183319P.  
PR 16-MAR-2000; 2000US-0189862P.  
PR 25-MAY-2000; 2000US-0207454P.  
PR 09-JUN-2000; 2000US-0211314P.  
PR 18-JUL-2000; 2000US-0219007P.  
PR 13-DEC-2000; 2000US-0255281P.  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
PI Schlegel R, Endege WO, Monahan JE;  
XX  
XX WPI; 2001-662795/76.  
DR  
XX  
PT Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer.  
XX  
XX Claim 1; Page 4469; 11750pp; English.  
XX  
XX The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for: (a) assessing whether  
CC a patient is afflicted with prostate cancer; (b) monitoring the  
CC progression of prostate cancer in a patient; (c) assessing the efficacy  
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
CC determining whether prostate cancer has metastasized in a patient; (h)  
CC assessing the aggressiveness or indolence of prostate cancer in a patient  
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX  
XX Sequence 1022 BP; 288 A; 203 C; 220 G; 311 T; 0 U; 0 Other;  
SQ  
Alignment Scores:  
Pred. No.: 5,29e-24 Length: 1022  
Score: 240.00 Matches: 47  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 97.92% Mismatches: 0  
Query Match: 78.18% Indels: 0  
DB: Gaps: 0  
US-10-031-158-14 (1-58) X ABV24072 (1-1022)  
QY 1 MetGIMetPheProSerProLeuPhePheLeuGlnLeuLeuYgInSerSer 20  
DB 589 ATGCAATGTTTCCCAAGCCCACTATTCTTCTTCAATTGCTGAACCAAGCTCC 530  
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40  
DB 529 AGAAGGCTGGAACATACCTTGTCTTGAGAAATTTTCCCTGATGTTATTAAAGATAC 470  
QY 41 IlegIlyblyblyArgArgAlaThr 48  
DB 469 ATTGGAGAGAAAGAAAGAGCATCG 446  
RESULT 36  
ABV24124/C  
ID ABV24124 standard; cDNA; 1022 BP.  
XX  
XX ABV24124;  
AC  
XX 16-SBP-2002 (first entry)  
DT Human prostate expression marker cDNA 24115.  
XX  
XX Human prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KM pharmacogenomic marker; gene; ss.  
XX

XX  
OS Homo sapiens.  
XX  
XX WO200160860-A2.  
PN  
XX  
XX 23-AUG-2001.  
PD  
XX  
PF 20-FEB-2001; 2001WO-US005171.  
XX  
PR 17-FEB-2000; 2000US-0183319P.  
PR 16-MAR-2000; 2000US-0189862P.  
PR 25-MAY-2000; 2000US-0207454P.  
PR 09-JUN-2000; 2000US-0211314P.  
PR 18-JUL-2000; 2000US-0219007P.  
PR 13-DEC-2000; 2000US-0255281P.  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
PI Schlegel R, Endege WO, Monahan JE;  
XX  
XX WPI; 2001-662795/76.  
DR  
XX  
PT Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer.  
XX  
XX Claim 1; Page 4484-4485; 11750pp; English.  
XX  
XX The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for: (a) assessing whether  
CC a patient is afflicted with prostate cancer; (b) monitoring the  
CC progression of prostate cancer in a patient; (c) assessing the efficacy  
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
CC determining whether prostate cancer has metastasized in a patient; (h)  
CC assessing the aggressiveness or indolence of prostate cancer in a patient  
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX  
XX Sequence 1022 BP; 288 A; 203 C; 220 G; 311 T; 0 U; 0 Other;  
SQ  
Alignment Scores:  
Pred. No.: 5,29e-24 Length: 1022  
Score: 240.00 Matches: 47  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 97.92% Mismatches: 0  
Query Match: 78.18% Indels: 0  
DB: Gaps: 0  
US-10-031-158-14 (1-58) X ABV24124 (1-1022)  
QY 1 MetGIMetPheProSerProLeuPhePheLeuGlnLeuLeuYgInSerSer 20  
DB 589 ATGCAATGTTTCCCAAGCCCACTATTCTTCTTCAATTGCTGAACCAAGCTCC 530  
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40  
DB 529 AGAAGGCTGGAACATACCTTGTCTTGAGAAATTTTCCCTGATGTTATTAAAGATAC 470  
QY 41 IlegIlyblyblyArgArgAlaThr 48  
DB 469 ATTGGAGAGAAAGAAAGAGCATCG 446  
RESULT 37  
ABV25823/C  
ID ABV25823 standard; cDNA; 1022 BP.  
XX  
XX ABV25823;  
AC  
XX 16-SBP-2002 (first entry)  
DT Human prostate expression marker cDNA 24115.  
XX  
XX Human prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KM pharmacogenomic marker; gene; ss.  
XX

DE Human prostate expression marker CDNA 25814.  
XX  
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
XX  
XX Homo sapiens.  
OS  
PN WO200160860-A2.  
XX  
XX 23-AUG-2001.  
PD  
PF 20-FEB-2001; 2001WO-US005171.  
XX  
XX 17-FEB-2000; 2000US-0183319P.  
PR 16-MAR-2000; 2000US-0189862P.  
PR 25-MAY-2000; 2000US-0207454P.  
PR 09-JUN-2000; 2000US-0211314P.  
PR 18-JUL-2000; 2000US-0219007P.  
PR 13-DEC-2000; 2000US-0255281P.  
XX  
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
PA  
PI Schlegel R, Endege WO, Monahan JB;  
XX  
XX WPI; 2001-662795/76.  
XX  
XX Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer.  
XX  
XX Claim 1; Page 5198; 11750pp; English.  
XX  
XX The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for: (a) assessing whether  
CC a patient is afflicted with prostate cancer; (b) monitoring the  
CC progression of prostate cancer in a patient; (c) assessing the efficacy  
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
CC determining whether prostate cancer has metastasized in a patient; (h)  
CC assessing the aggressiveness or indolence of prostate cancer in a patient  
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX  
XX SQ Sequence 1022 BP; 288 A; 203 C; 220 G; 311 T; 0 U; 0 Other;  
XX  
XX Alignment Scores:  
Pred. No.: 5,29e-24 Length: 1022  
Score: 240.00 Matches: 47  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 97.92% Mismatches: 0  
Query Match: 78.18% Indels: 0  
Gaps: 0  
DB: 5  
US-10-031-158-14 (1-58) x ABV25823 (1-1022)  
QY 1 MetGlnMetPheProPheProSerProLeuPhePhePheLeuGlnLeuLeuYgInSerSer 20  
Db 589 ATGCAAGATGTTTCCCAAGCCCACTATTTTCTTCAATTGCTGAAACAAAGCTCC 530  
QY 21 ArgArgLeuGlnIuhIsthrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40  
Db 529 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAATAC 470  
QY 41 IleGlyIyelysArgArgAlaThr 48  
Db 469 ATTGGGAAGAAAAGAGCATCG 446  
RESULT 38  
ABV24034/C  
ID ABV24034 standard; cDNA; 1022 BP.  
XX

AC ABV24034;  
XX  
XX 16-SEP-2002 (first entry)  
XX  
XX Human prostate expression marker CDNA 24025.  
DE  
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
XX  
XX Homo sapiens.  
OS  
PN WO200160860-A2.  
XX  
XX 23-AUG-2001.  
PD  
PF 20-FEB-2001; 2001WO-US005171.  
XX  
XX 17-FEB-2000; 2000US-0183319P.  
PR 16-MAR-2000; 2000US-0189862P.  
PR 25-MAY-2000; 2000US-0207454P.  
PR 09-JUN-2000; 2000US-0211314P.  
PR 18-JUL-2000; 2000US-0219007P.  
PR 13-DEC-2000; 2000US-0255281P.  
XX  
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
PA  
PI Schlegel R, Endege WO, Monahan JB;  
XX  
XX WPI; 2001-662795/76.  
XX  
XX Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer.  
XX  
XX Claim 1; Page 4456; 11750pp; English.  
XX  
XX The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for: (a) assessing whether  
CC a patient is afflicted with prostate cancer; (b) monitoring the  
CC progression of prostate cancer in a patient; (c) assessing the efficacy  
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
CC determining whether prostate cancer has metastasized in a patient; (h)  
CC assessing the aggressiveness or indolence of prostate cancer in a patient  
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX  
XX SQ Sequence 1022 BP; 288 A; 203 C; 220 G; 311 T; 0 U; 0 Other;  
XX  
XX Alignment Scores:  
Pred. No.: 5,29e-24 Length: 1022  
Score: 240.00 Matches: 47  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 97.92% Mismatches: 0  
Query Match: 78.18% Indels: 0  
Gaps: 0  
DB: 5  
US-10-031-158-14 (1-58) x ABV24034 (1-1022)  
QY 1 MetGlnMetPheProPheProSerProLeuPhePhePheLeuGlnLeuLeuYgInSerSer 20  
Db 589 ATGCAAGATGTTTCCCAAGCCCACTATTTTCTTCAATTGCTGAAACAAAGCTCC 530  
QY 21 ArgArgLeuGlnIuhIsthrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40  
Db 529 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAATAC 470  
QY 41 IleGlyIyelysArgArgAlaThr 48  
Db 469 ATTGGGAAGAAAAGAGCATCG 446

RESULT 39  
ABV23986/c  
ID ABV23986 standard; cDNA; 1022 BP.  
XX  
AC ABV23986;  
XX  
DT 16-SEP-2002 (first entry)  
XX  
DE Human prostate expression marker cDNA 23977.  
XX  
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN W0200160860-A2.  
XX  
PD 23-AUG-2001.  
XX  
PF 20-FEB-2001; 2001WO-US005171.  
XX  
PR 17-FEB-2000; 2000US-0183319P.  
PR 16-MAR-2000; 2000US-0189862P.  
PR 25-MAY-2000; 2000US-0207454P.  
PR 09-JUN-2000; 2000US-0211314P.  
PR 18-JUL-2000; 2000US-0219007P.  
PR 13-DEC-2000; 2000US-0255281P.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
PI Schlegel R, Endege WO, Monahan JE;  
XX  
DR WPI; 2001-662795/76.  
XX  
PT Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer.  
XX  
PS Claim 1; Page 4439; 11750pp; English.  
XX  
CC The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for: (a) assessing whether  
CC a patient is afflicted with prostate cancer; (b) monitoring the  
CC progression of prostate cancer in a patient; (c) assessing the efficacy  
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
CC determining whether prostate cancer has metastasized in a patient; (h)  
CC assessing the aggressiveness or indolence of prostate cancer in a patient  
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX  
SQ Sequence 1022 BP; 288 A; 203 C; 220 G; 311 T; 0 U; 0 Other;  
XX  
Alignment Scores:  
Pred. No.: 5,29e-24 Length: 1022  
Score: 240.00 Matches: 47  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 97.92% Mismatches: 0  
Query Match: 78.18% Indels: 0  
DB: Gaps: 0  
US-10-031-158-14 (1-58) x ABV23986 (1-1022)  
QY 1 MetGlnMetPheProSerProLeuPhePhePheLeuGlnLeuLeuYsgInSerSer 20  
DB 589 ATGCAATGTTTCCCAAGCCCACTATTTTCTTCTCAATGCTGAAACAAAGCTCC 530  
QY 21 ArgArgLeuGlnIuhIsthrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40  
DB 529 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAATAC 470

QY 41 lIegIyLySLyArgArgAlaThr 48  
DB 469 ATTGGAGAGAAAAGAGAGCATCG 446  
RESULT 40  
ABV23991/c  
ID ABV23991 standard; cDNA; 1022 BP.  
XX  
AC ABV23991;  
XX  
DT 16-SEP-2002 (first entry)  
XX  
DE Human prostate expression marker cDNA 23982.  
XX  
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN W0200160860-A2.  
XX  
PD 23-AUG-2001.  
XX  
PF 20-FEB-2001; 2001WO-US005171.  
XX  
PR 17-FEB-2000; 2000US-0183319P.  
PR 16-MAR-2000; 2000US-0189862P.  
PR 25-MAY-2000; 2000US-0207454P.  
PR 09-JUN-2000; 2000US-0211314P.  
PR 18-JUL-2000; 2000US-0219007P.  
PR 13-DEC-2000; 2000US-0255281P.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
PI Schlegel R, Endege WO, Monahan JE;  
XX  
DR WPI; 2001-662795/76.  
XX  
PT Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer.  
XX  
PS Claim 1; Page 4442; 11750pp; English.  
XX  
CC The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for: (a) assessing whether  
CC a patient is afflicted with prostate cancer; (b) monitoring the  
CC progression of prostate cancer in a patient; (c) assessing the efficacy  
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing  
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)  
CC determining whether prostate cancer has metastasized in a patient; (h)  
CC assessing the aggressiveness or indolence of prostate cancer in a patient  
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker  
XX  
SQ Sequence 1022 BP; 288 A; 203 C; 220 G; 311 T; 0 U; 0 Other;  
XX  
Alignment Scores:  
Pred. No.: 5,29e-24 Length: 1022  
Score: 240.00 Matches: 47  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 97.92% Mismatches: 0  
Query Match: 78.18% Indels: 0  
DB: Gaps: 0  
US-10-031-158-14 (1-58) x ABV23991 (1-1022)  
QY 1 MetGlnMetPheProSerProLeuPhePhePheLeuGlnLeuLeuYsgInSerSer 20  
DB 589 ATGCAATGTTTCCCAAGCCCACTATTTTCTTCTCAATGCTGAAACAAAGCTCC 530

Qy	21	ArgArgLeuGluHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr	40
Db	529	AGAAAGGCTGGAACATACCTTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAAGATAC	470
Qy	41	ILEGlyLysLysArgArgAlaThr	48
Db	469	ATTGGAGAGAAAGAAAGCATCG	446

Search completed: December 8, 2004, 08:22:18  
Job time : 253 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 8, 2004, 07:44:01 ; Search time 1517 Seconds  
(without alignments)  
1393.213 Million cell updates/sec

Title: US-10-031-158-14  
Perfect score: 307  
Sequence: 1 MQMPPSPPLFFFLQQLKQSS.....RYIGKKRRARFMDPRRGTG 58

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2.1/USPRO.spool/US10031158/runat.06122004.082659.15900/app\_query.fasta\_1.199  
-DB=EST -OPMT=fastap -SUFFIX=ret -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=40 -MODE=LOCAL  
-OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10031158 @CGN 1.1 2607 @runat.06122004.082659.15900 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -MANY\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
EST:  
1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_hc.\*  
4: gb\_est3.\*  
5: gb\_est4.\*  
6: gb\_est5.\*  
7: gb\_est6.\*  
8: gb\_g881.\*  
9: gb\_g882.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	307	100.0	427	2	BF679325 602153825
2	307	100.0	510	1	AV716204 AV716204
3	307	100.0	596	1	AV715641 AV715641
4	307	100.0	603	1	AV714189 AV714189
5	307	100.0	803	1	AV756492 AV756492
6	307	100.0	818	4	BI771070 603059635
7	307	100.0	821	2	BF679123 602153390
8	307	100.0	861	2	BF681385 602156579
9	307	100.0	865	1	AI557112 PT2.1_13_

10	307	100.0	881	2	BF677648	BF677648 602085529
11	307	100.0	902	2	BF674593	BF674593 602137811
12	307	100.0	927	3	BF679476	BF679476 602153280
13	307	100.0	1058	4	BM544213	BM544213 AGENCOURT
14	307	100.0	3533	3	BC030554	BC030554 Homo sapi
15	304	99.0	789	5	BU199143	BU199143 DCBCL108
16	294	95.8	653	2	BE974171	BE974171 601680522
17	294	95.8	1005	4	BM919380	BM919380 AGENCOURT
18	294	95.8	1757	3	BC039116	BC039116 Homo sapi
19	290	94.5	434	2	BF747659	BF747659 OVE-BT061
20	290	94.5	477	2	AM087486	AM087486 XB57C01.X
21	290	94.5	681	1	AV714756	AV714756 AV714756
22	290	94.5	1510	3	BC017861	BC017861 Homo sapi
23	289	94.1	1055	2	BF964904	BF964904 602268596
24	288	93.8	523	2	BF675605	BF675605 602083495
25	284	92.5	362	1	AI758378	AI758378 tV68C12.X
26	278	90.6	719	2	BF681238	BF681238 602155502
27	270	87.9	854	2	BF681029	BF681029 602155433
28	268	87.3	629	2	BF679165	BF679165 602153737
29	267	87.0	803	2	BF680840	BF680840 602156088
30	262	85.3	772	4	BI836020	BI836020 603084004
31	259	84.4	808	2	BF673289	BF673289 602136139
32	253	82.4	857	2	BF674457	BF674457 602137231
33	251	81.8	849	2	BF679284	BF679284 602153477
34	248	80.8	995	7	CO647648	CO647648 ILLUMIGEN
35	245	79.8	663	6	CB553506	CB553506 MMSPP004
36	227	73.9	870	2	BF678129	BF678129 602085181
37	226	73.6	637	6	CB551454	CB551454 MMSPP013
38	224	73.0	670	2	BF677916	BF677916 602084736
39	221	72.0	827	4	BG535408	BG535408 602563069
40	209	68.1	572	2	BF370026	BF370026 MR3-FN000
41	204.5	66.6	864	4	BI834993	BI834993 603088885
42	203	66.1	552	4	BP164784	BP164784 BP164784
43	200	65.1	572	5	BP162849	BP162849 BP162849
44	200	65.1	888	5	BP162346	BP162346 BP162346
45	193	62.9	436	5	BP165939	BP165939 BP165939

## ALIGNMENTS

RESULT 1  
BF679325  
LOCUS  
DEFINITION  
602153825F1 NIH\_MGC\_83 Homo sapiens cDNA clone IMAGE:4294566 5',  
mRNA sequence.

ACCESSION  
BF679325  
VERSION  
BF679325.1 GI:11953220

KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)

ORGANISM  
Homo sapiens

REFERENCE  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 427)

AUTHORS  
NIH-MGC <http://mgc.ncl.nih.gov/>.  
TITLE  
National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL  
Unpublished (1999)  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)

REFERENCE  
Tissue Procurement: CLONETECH Laboratories, Inc.  
CDNA Library Preparation: CLONETECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L1CM1144 row: b column: 07  
High quality sequence stop: 427.

## FEATURES

source  
location/Qualifiers  
1..427  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4294566"

/lab host="DH10B (T1 phage-resistant)"  
/clone lib="NH MGC\_83"  
/note="Organ: prostate; Vector: pDNR-LIB (Clontech);  
Site 1: Sfil (ggcgctcgcc); Site 2: Sfil  
(ggcgctcgcc); 5' and 3' adaptors were used in cloning  
as follows: 5' adaptor sequence: 5'-CACGCGCATTTAGGCC-3'  
and 3' adaptor sequence:  
5'-ATCTAGAGCGCGCGCGCCGACATG-dt(30)BN-3' (where B = A,  
C, or G and N = A, C, or T). Average insert size 1.4  
kb (range 0.5-4.0 kb). 14/15 colonies contained inserts  
by PCR. This library was enriched for full-length clones  
and was constructed by Clontech Laboratories (Palo Alto,  
CA)."

ORIGIN

Alignment Scores:

Pred. No.:	1,17e-30	Length:	427
Score:	307.00	Matches:	58
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-10-031-158-14 (1-58) x BF679325 (1-427)

Qy 1 MetGlnMetPheProSerProLeuPhePhePheLeuGlnLeuLeuLeuGlnSerSer 20  
Db 66 ATGCAGATGTTTCCCGCAAGCCCACTATTTTCTTCAATTCGTAACAAAGCTCC 125

Qy 21 ArgArgLeuGlnUhiThrPheValPheLeuArganPheSerLeuMetLeuLeuArgTyr 40  
Db 126 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAGATAC 185

Qy 41 lIleGlyLysIleArgArgAlaThrArgPheTrpPheProArgArgGlyThrPro 58  
Db 186 ATTGGCAAGAAAGAAAGACACACGATTTGGATCCAGAGGAGAACACCA 239

RESULT 2  
LOCUS AV716204 510 bp mRNA linear EST 11-OCT-2000  
DEFINITION AV716204 DCB Homo sapiens cDNA clone DCBAM08 5', mRNA sequence.  
ACCESSION AV716204  
VERSION AV716204.1 GI:10797721  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 510)  
Xu,X., Gu,J., Liu,F., Qu,J., Zhao,M., Li,Y., Huang,Q., Zhou,J.,  
Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X.,  
Cheng,Z., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M.,  
Lu,G., Cheng,Z. and Han,Z.  
Homo sapiens cDNA DCB clones  
Unpublished (2000)  
Contact: Zeguang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
201203, P. R. China  
Tel: 86-21-50801919(ex.45)  
Fax: 86-21-50801922  
Email: hanzg@chgc.sh.cn  
This clone is available at CHGC in Shanghai.  
Location/Qualifiers

FEATURES  
source  
1..510  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DCBAM08"  
/cell\_type="dendritic cells"  
/dev\_stage="mature"  
/lab\_host="BM25.8"  
/clone\_lib="DCB"

/note="Vector: pTriplEx2; Site\_1: SfilA; Site\_2: SfilB"

ORIGIN

Alignment Scores:

Pred. No.:	1,45e-30	Length:	510
Score:	307.00	Matches:	58
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0

US-10-031-158-14 (1-58) x AV716204 (1-510)

Qy 1 MetGlnMetPheProSerProLeuPhePhePheLeuGlnLeuLeuLeuGlnSerSer 20  
Db 159 ATGCAGATGTTTCCCGCAAGCCCACTATTTTCTTCAATTCGTAACAAAGCTCC 218

Qy 21 ArgArgLeuGlnUhiThrPheValPheLeuArganPheSerLeuMetLeuLeuArgTyr 40  
Db 219 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAGATAC 278

Qy 41 lIleGlyLysIleArgArgAlaThrArgPheTrpPheProArgArgGlyThrPro 58  
Db 279 ATTGGCAAGAAAGAAAGACACGATTTGGATCCAGAGGAGAACACCA 332

RESULT 3  
LOCUS AV715641 596 bp mRNA linear EST 11-OCT-2000  
DEFINITION AV715641 DCB Homo sapiens cDNA clone DCBBS05 5', mRNA sequence.  
ACCESSION AV715641  
VERSION AV715641.1 GI:10797158  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 596)  
Xu,X., Gu,J., Liu,F., Qu,J., Zhao,M., Li,Y., Huang,Q., Zhou,J.,  
Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X.,  
Cheng,Z., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M.,  
Lu,G., Cheng,Z. and Han,Z.  
Homo sapiens cDNA DCB clones  
Unpublished (2000)  
Contact: Zeguang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
201203, P. R. China  
Tel: 86-21-50801919(ex.45)  
Fax: 86-21-50801922  
Email: hanzg@chgc.sh.cn  
This clone is available at CHGC in Shanghai.  
Location/Qualifiers

FEATURES  
source  
1..596  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DCBBS05"  
/cell\_type="dendritic cells"  
/dev\_stage="mature"  
/lab\_host="BM25.8"  
/clone\_lib="DCB"  
/note="Vector: pTriplEx2; Site\_1: SfilA; Site\_2: SfilB"

ORIGIN

Alignment Scores:

Pred. No.:	1,74e-30	Length:	596
Score:	307.00	Matches:	58
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0

US-10-031-158-14 (1-58) x AV715641 (1-596)

QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuLyGlnSerSer 20  
 DB 93 ATCAAGATGTTTCCCCCAAGCCCACTATTTTCTTCTCAATTGCTGAACAAAGCTCC 152  
 QY 21 ArgArgLeuGlnIuIsthrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40  
 DB 153 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAAGATAC 212  
 QY 41 IleglYlylYsArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58  
 DB 213 ATTGGCAAGAAAAGAGACCAACGATTCGTGGATCCCGAGGGGGAACACCA 266  
 RESULT 4  
 AV714189 603 bp mRNA linear EST 11-OCT-2000  
 LOCUS AV714189 DCB Homo sapiens cDNA clone DCB0E09 5', mRNA sequence.  
 ACCESSION AV714189  
 VERSION AV714189.1 GI:10795706  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 COMMENT Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 603)  
 AUTHORS Xu,X., Gu,J., Liu,F., Qu,J., Zhao,M., Li,Y., Huang,Q., Zhou,J., Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X., Cheng,Z., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Cheng,Z. and Han,Z.  
 TITLE Homo sapiens CDNA DCB clones  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Zeguang Han  
 Chinese National Human Genome Center at Shanghai  
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China  
 Tel: 86-21-50801919 (ex.45)  
 Fax: 86-21-50801922  
 Email: hanzg@chgc.sh.cn  
 This clone is available at CHGC in Shanghai.  
 FEATURES  
 source  
 1. 603  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="DCB0E09"  
 /cell\_type="dendritic cells"  
 /dev\_stage="mature"  
 /lab\_host="BM25.8"  
 /clone\_lib="DCB"  
 /note="Vector: pTriplEx2; Site\_1: sfilA; Site\_2: sfilB"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.77e-30 Length: 603  
 Score: 307.00 Matches: 58  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 1 Gaps: 0  
 US-10-031-158-14 (1-58) x AV714189 (1-603)  
 QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuLyGlnSerSer 20  
 DB 252 ATGAGATGTTTCCCCCAAGCCCACTATTTTCTTCTCAATTGCTGAACAAAGCTCC 311  
 QY 21 ArgArgLeuGlnIuIsthrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40  
 DB 312 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAAGATAC 371  
 QY 41 IleglYlylYsArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58  
 DB 372 ATTGGCAAGAAAAGAGACCAACGATTCGTGGATCCCGAGGGGGAACACCA 425

RESULT 5  
 AV756492 803 bp mRNA linear EST 19-OCT-2000  
 LOCUS AV756492 BM Homo sapiens cDNA clone BMFBAF12 5', mRNA sequence.  
 ACCESSION AV756492  
 VERSION AV756492.1 GI:10914340  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 COMMENT Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 803)  
 AUTHORS Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H., Gu,Y., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Yang,Y., Gao,G., Wang,Z., Zhang,Q., Chen,S., Han,Z. and Chen,Z.  
 TITLE Homo sapiens CDNA BM clones  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Zeguang Han  
 Chinese National Human Genome Center at Shanghai  
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China  
 Tel: 86-21-50801919 (ex.45)  
 Fax: 86-21-50801922  
 Email: hanzg@chgc.sh.cn  
 This clone is available at CHGC in Shanghai.  
 FEATURES  
 source  
 1. 803  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="BMFBAF12"  
 /tissue="Bone marrow"  
 /cell\_type="CD34+ hematopoietic stem/progenitor cell"  
 /lab\_host="BM25.8"  
 /clone\_lib="BM"  
 /note="Vector: pTriplEx2; Site\_1: sfilA; Site\_2: sfilB"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2.48e-30 Length: 803  
 Score: 307.00 Matches: 58  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 1 Gaps: 0  
 US-10-031-158-14 (1-58) x AV756492 (1-803)  
 QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuLyGlnSerSer 20  
 DB 434 ATCAAGATGTTTCCCCCAAGCCCACTATTTTCTTCAATTGCTGAACAAAGCTCC 493  
 QY 21 ArgArgLeuGlnIuIsthrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40  
 DB 494 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAAGATAC 553  
 QY 41 IleglYlylYsArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58  
 DB 554 ATTGGCAAGAAAAGAGACCAACGATTCGTGGATCCCGAGGGGGAACACCA 607  
 RESULT 6  
 BI771070 818 bp mRNA linear EST 25-SEP-2001  
 LOCUS BI771070 603059635F1 NIH\_MGC\_122 Homo sapiens cDNA clone IMAGE:5208945 5', mRNA sequence.  
 ACCESSION BI771070  
 VERSION BI771070.1 GI:15762648  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 818)  
 AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 JOURNAL NIH-MGC http://mgc.nci.nih.gov/  
 COMMENT National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)  
 CONTACT: Robert Strausberg, Ph.D.  
 EMAIL: cgsbds-remail.nih.gov  
 TISSUE Procurement: Life Technologies, Inc.  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov  
 Plate: LHAM1524 row: e column: 10  
 High quality sequence stop: 816.  
 Location/Qualifiers  
 1. 818  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5208945"  
 /lab\_host="DH10B"  
 /clone\_1id="NIH\_MGC\_122"  
 /note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung; 16 week female spleen, and 20-22 week male spleen. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH\_MGC Library."  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2,546-30 Length: 818  
 Score: 307.00 Matches: 58  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0  
 US-10-031-158-14 (1-58) x B1771070 (1-818)  
 QY 1 MetGlnMetPheProPheProLeuPhePheLeuGlnLeuLeuGlnSerSer 20  
 Db 548 ATGCAGATGTTTCCCAAGCCCACTATTTCTTCCTCAATTGCTGAACAAAGCTCC 607  
 QY 21 ArgArgLeuGluHisThrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40  
 Db 608 AGAAGGCTGGAACATACCTTGTCTCTTGAGAAATTTTCCCTGATGTTATTAGATAC 667  
 QY 41 IlleGlyLysLysArgArgAlaThrArgPheThrPheProArgArgGlyThrPro 58  
 Db 668 ATTGGCAAGAAAGAGAGCAACAGATTCTGGATCCAGAGAGGGAACACCA 721  
 RESULT 7  
 LOCUS BF679123 821 bp mRNA linear EST 21-DEC-2000  
 DEFINITION 602153390F1 NIH\_MGC\_83 Homo sapiens cDNA clone IMAGE:4294247 5', mRNA sequence.  
 ACCESSION BF679123  
 VERSION BF679123.1 GI:11953018  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 821)  
 JOURNAL NIH-MGC http://mgc.nci.nih.gov/  
 COMMENT

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 EMAIL: cgsbds-remail.nih.gov  
 TISSUE Procurement: CLONTECH Laboratories, Inc.  
 CDNA Library Preparation: CLONTECH Laboratories, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov  
 Plate: LHCMI143 row: d column: 24  
 High quality sequence stop: 620.  
 Location/Qualifiers  
 1. 821  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4294247"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_1id="NIH\_MGC\_83"  
 /note="Organ: prostate; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggcgccctggcc); Site 2: SfiI (ggcgatcatggcc); 5' and 3' adaptor were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGACC-3' and 3' adaptor sequence: 5'-ATTCTAAGCGCGGCGGCACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.4 kb (range 0.5-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2,556-30 Length: 821  
 Score: 307.00 Matches: 58  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 2 Gaps: 0  
 US-10-031-158-14 (1-58) x BF679123 (1-821)  
 QY 1 MetGlnMetPheProPheProLeuPhePheLeuGlnLeuLeuGlnSerSer 20  
 Db 64 ATGCAGATGTTTCCCAAGCCCACTATTTCTTCCTCAATTGCTGAACAAAGCTCC 123  
 QY 21 ArgArgLeuGluHisThrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40  
 Db 124 AGAAGGCTGGAACATACCTTGTCTCTTGAGAAATTTTCCCTGATGTTATTAGATAC 183  
 QY 41 IlleGlyLysLysArgArgAlaThrArgPheThrPheProArgArgGlyThrPro 58  
 Db 184 ATTGGCAAGAAAGAGAGCAACAGATTCTGGATCCAGAGAGGGAACACCA 237  
 RESULT 8  
 LOCUS BF681385 861 bp mRNA linear EST 21-DEC-2000  
 DEFINITION 602156579F1 NIH\_MGC\_83 Homo sapiens cDNA clone IMAGE:4297555 5', mRNA sequence.  
 ACCESSION BF681385  
 VERSION BF681385.1 GI:11955280  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 861)  
 JOURNAL NIH-MGC http://mgc.nci.nih.gov/  
 COMMENT

Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: CLONTECH Laboratories, Inc.  
 CDNA Library Preparation: CLONTECH Laboratories, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
 Plate: LNCM151 row: n column: 20  
 High quality sequence stop: 672.

## FEATURES

source

Location/Qualifiers

1..861

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone\_image="IMAGE:429755"

/lab\_host="DH10B (TI phage-resistant)"

/clone\_lib="NIH\_MGC\_83"

/note="Organ: prostate; Vector: pDNR-LIB (Clontech); Site: 1: SfiI (ggcgctgggc); Site 2: SfiI (ggcgctgggc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCCGCGCCGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.4 kb (range 0.5-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

## ORIGIN

## Alignment Scores:

Pred. No.:	2.7e-30	Length:	861
Score:	307.00	Matches:	58
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-10-031-158-14 (1-58) x BF681385 (1-861)

Qy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20  
 Db 69 ATGCAGATGTTTCCCAAGCCCACTATTTTCTTCTTCATTCGCGAACAAGCTCC 128  
 Qy 21 ArgArgLeuGlnHisThrPheValPheLeuArgAspPheSerLeuMetLeuLeuArgTyr 40  
 Db 129 AGAAGCTGGAACATACCTTGTCTTGTGAGAAATTTTCCCTGATGTTATTAAATAC 188  
 Qy 41 TlGlyLysLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58  
 Db 189 ATTGGCAAGAAAGAGAGCAACGATTCGGATCCAGAGGGAGCAACCA 242

RESULT 9  
 A1557112 865 bp mRNA linear EST 09-AUG-1999  
 LOCUS PT2.1.13.F05.r tumor2 Homo sapiens CDNA 3', mRNA sequence.  
 ACCESSION A1557112  
 VERSION A1557112.1 GI:4489475  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 865)  
 Huang, G.M., Ng, W.L., Parkes, J., He, L., Liang, H.A., Gordon, D., Yu, J., and Hood, L.  
 Prostate cancer expression profiling by cDNA sequencing analysis  
 JOURNAL Genomics 59 (2), 178-186 (1999)  
 MEDLINE 99339982  
 PUBMED 10409429  
 COMMENT Leroy Hood

University of Washington  
 Department of Molecular Biotechnology, Box 357730, University of  
 Washington, Seattle, WA 98195  
 Tel: 5106280100  
 Fax: 5106280108  
 Email: huangm@yahoo.com.

## FEATURES

source

Location/Qualifiers

1..865

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone\_lib="tumor2"

/note="Organ: Prostate; Vector: pBluescript; Directional CDNA library was constructed using Lambda ZP II kit (Stratagene). mRNA was extracted from a frozen prostate tumor tissue (Mayo Clinics)."

## ORIGIN

## Alignment Scores:

Pred. No.:	2.71e-30	Length:	865
Score:	307.00	Matches:	58
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0

US-10-031-158-14 (1-58) x A1557112 (1-865)

Qy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20  
 Db 59 ATGCAGATGTTTCCCAAGCCCACTATTTTCTTCTTCATTCGCGAACAAGCTCC 118  
 Qy 21 ArgArgLeuGlnHisThrPheValPheLeuArgAspPheSerLeuMetLeuLeuArgTyr 40  
 Db 119 AGAAGCTGGAACATACCTTGTCTTGTGAGAAATTTTCCCTGATGTTATTAAATAC 178  
 Qy 41 TlGlyLysLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58  
 Db 179 ATTGGCAAGAAAGAGAGCAACGATTCGGATCCAGAGGGAGCAACCA 232

RESULT 10  
 BF677648 881 bp mRNA linear EST 21-DEC-2000  
 LOCUS BF677648  
 DEFINITION 602085529F1 NIH\_MGC\_83 Homo sapiens CDNA clone IMAGE:4249807 5', mRNA sequence.  
 ACCESSION BF677648  
 VERSION BF677648.1 GI:11951543  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 881)  
 NIH-MGC <http://imgc.ncbi.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Straube, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: CLONTECH Laboratories, Inc.  
 CDNA Library Preparation: CLONTECH Laboratories, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
 Plate: LNCM1070 row: 1 column: 08  
 High quality sequence stop: 601.

FEATURES  
 source  
 Location/Qualifiers  
 1..881  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone\_image="IMAGE:4249807"

/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NIH\_MGC\_83"  
/note="Organ: prostate; Vector: pDNR-LIB (Clontech);  
Site 1: SfiI (ggcgctcgcc); Site 2: SfiI  
(ggcgctcgcc); 5' and 3' adaptor were used in cloning  
as follows: 5'-CACGGCCATTATGGCC-3'  
and 3' adaptor sequence:  
5'-ATTCTAGAGCCGAGCGCCGACATG-dt(30)BN-3' (where B = A,  
C, or G and N = A, C, G, or T). Average insert size 1.4  
kb (range 0.5-4.0 kb). 14/15 colonies contained inserts  
by PCR. This library was enriched for full-length clones  
and was constructed by Clontech Laboratories (Palo Alto,  
CA)."

ORIGIN

Alignment Scores:

Pred. No.:	2.77e-30	Length:	881
Score:	307.00	Matches:	58
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-10-031-158-14 (1-58) x BF677648 (1-881)

Qy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20  
Db 64 ATGCAGATGTTTCCCAAGCCCACTATTTCTTCTTCATTGCTGAACAAAGCTCC 123

Qy 21 ArgArgLeuGlnHisThrPheValPheLeuArgAnpSerLeuMetLeuLeuArgTyr 40  
Db 124 AGAAGGCTGGAACATACCTTGCTCTTGAGAAATTTTCCCTGATGTATTAAGATAC 183

Qy 41 IleglylylsySAAGAGAlaThrArgPheTrpAspProArgArgGlyThrPro 58  
Db 184 ATTGGGAAGAAAGAGAGACAGATTCGTGGATCCAGAGAGGGAACACCA 237

RESULT 11  
LOCUS BF674593 902 bp mRNA linear EST 21-DEC-2000  
DEFINITION 60213781P1 NIH\_MGC\_83 Homo sapiens cDNA clone IMAGE:4274283 5',  
mRNA sequence.  
ACCESSION BF674593  
VERSION BF674593.1 GI:11948488  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.  
AUTHORS 1 (bases 1 to 902)  
TITLE NIH-MGC  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@bde-remail.nih.gov  
Tissue Procurement: CLONETECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
http://image.llnl.gov  
Plate: L1CM1091 row: e column: 04  
High quality sequence stop: 629.  
Location/Qualifiers  
1..902  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4274283"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NIH\_MGC\_83"  
/note="Organ: prostate; Vector: pDNR-LIB (Clontech);  
Site 1: SfiI (ggcgctcgcc); Site 2: SfiI  
(ggcgctcgcc); 5' and 3' adaptor were used in cloning  
as follows: 5'-CACGGCCATTATGGCC-3'

FEATURES  
source

Site 1: SfiI (ggcgctcgcc); Site 2: SfiI  
(ggcgctcgcc); 5' and 3' adaptor were used in cloning  
as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3'  
and 3' adaptor sequence:  
5'-ATTCTAGAGCCGAGCGCCGACATG-dt(30)BN-3' (where B = A,  
C, or G and N = A, C, G, or T). Average insert size 1.4  
kb (range 0.5-4.0 kb). 14/15 colonies contained inserts  
by PCR. This library was enriched for full-length clones  
and was constructed by Clontech Laboratories (Palo Alto,  
CA)."

ORIGIN

Alignment Scores:

Pred. No.:	2.85e-30	Length:	902
Score:	307.00	Matches:	58
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-10-031-158-14 (1-58) x BF674593 (1-902)

Qy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20  
Db 64 ATGCAGATGTTTCCCAAGCCCACTATTTCTTCTTCATTGCTGAACAAAGCTCC 123

Qy 21 ArgArgLeuGlnHisThrPheValPheLeuArgAnpSerLeuMetLeuLeuArgTyr 40  
Db 124 AGAAGGCTGGAACATACCTTGCTCTTGAGAAATTTTCCCTGATGTATTAAGATAC 183

Qy 41 IleglylylsySAAGAGAlaThrArgPheTrpAspProArgArgGlyThrPro 58  
Db 184 ATTGGGAAGAAAGAGAGACAGATTCGTGGATCCAGAGAGGGAACACCA 237

RESULT 12  
LOCUS BF679476 927 bp mRNA linear EST 21-DEC-2000  
DEFINITION 602153280P1 NIH\_MGC\_83 Homo sapiens cDNA clone IMAGE:4294507 5',  
mRNA sequence.  
ACCESSION BF679476  
VERSION BF679476.1 GI:11953360  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.  
AUTHORS 1 (bases 1 to 927)  
TITLE NIH-MGC  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@bde-remail.nih.gov  
Tissue Procurement: CLONETECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
http://image.llnl.gov  
Plate: L1CM1143 row: o column: 20  
High quality sequence stop: 573.  
Location/Qualifiers  
1..927  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4294507"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NIH\_MGC\_83"  
/note="Organ: prostate; Vector: pDNR-LIB (Clontech);  
Site 1: SfiI (ggcgctcgcc); Site 2: SfiI  
(ggcgctcgcc); 5' and 3' adaptor were used in cloning  
as follows: 5'-CACGGCCATTATGGCC-3'

FEATURES  
source

and 3' adaptor sequence:  
5'-ATTCTAGAGCGCGCGCCAGATG-dr(30)BN-3' (where B = A,  
C, or G and N = A, C, G, or T). Average insert size 1.4  
kb (range 0.5-4.0 kb). 14/15 colonies contained inserts  
by PCR. This library was enriched for full-length clones  
and was constructed by Clontech Laboratories (Palo Alto,  
CA)."

## ORIGIN

## Alignment Scores:

Pred. No.:	2.94e-30	Length:	927
Score:	307.00	Matches:	58
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-10-031-158-14 (1-58) x BF679476 (1-927)

Qy 1 MetGlnMetPheProPheProLeuPhePhePheLeuGlnLeuLeuGlnSerSer 20  
Db 67 ATGAGATGTTTCCCCCAAGCCCACTATTTTCTTCTCAATTGCTGAACAAGCTCC 126

Qy 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40  
Db 127 AGAAGCTGGAACATACCTTGTCTCTTGAGAAATTTTCCCTGATGTTATTAAAGTAC 186

Qy 41 TleGlyLysLysArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58  
Db 187 ATTGGCAAGAAAGAGAGACAGATTCGTGGATCCAGAGGGGAGAACCA 240

## RESULT 13

LOCUS BMS44213 1058 bp mRNA linear EST 20-FEB-2002  
DEFINITION AGENCOURT\_6490556 NIH\_MGC\_125 Homo sapiens cDNA clone IMAGE:5587705  
5', mRNA sequence.  
ACCESSION BMS44213  
VERSION BMS44213.1 GI:18775279  
KEYWORDS EST.

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 1058)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)  
Tissue Procurement: Invitrogen

CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>

Plate: LLM12357 Row: k Column: 02  
High quality sequence start: 12  
High quality sequence stop: 768.

## FEATURES

source

1..1058  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5587705"  
/lab\_host="DH10B"  
/clone\_id="NIH\_MGC\_125"  
/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;  
Site 1: EcoRV (destroyed); Site 2: NotI; RNA source pool  
of three ovaries, from females ranging in age from 38 to  
49 yo. Library is oligo-dT primed and directionally cloned  
(EcoRV site is destroyed upon cloning). Average insert  
size 2.1 kb, insert size range 1-3.5 kb. Library is

normalized and enriched for full-length clones and was  
constructed by C. Gruber (Invitrogen). Research Genetics  
tracking code 036."

## ORIGIN

Alignment Scores:	3.44e-30	Length:	1058
Pred. No.:	307.00	Matches:	58
Score:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-10-031-158-14 (1-58) x BMS44213 (1-1058)

Qy 1 MetGlnMetPheProPheProLeuPhePhePheLeuGlnLeuLeuGlnSerSer 20  
Db 421 ATGAGATGTTTCCCCCAAGCCCACTATTTTCTTCTCAATTGCTGAACAAGCTCC 480

Qy 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40  
Db 481 AGAAGCTGGAACATACCTTGTCTCTTGAGAAATTTTCCCTGATGTTATTAAAGTAC 540

Qy 41 TleGlyLysLysArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58  
Db 541 ATTGGCAAGAAAGAGAGACAGATTCGTGGATCCAGAGGGGAGAACCA 594

## RESULT 14

LOCUS BC030554 3533 bp mRNA linear HTC 20-MAY-2002  
DEFINITION Homo sapiens, similar to T cell receptor gamma locus, clone  
IMAGE:5213435, mRNA.  
ACCESSION BC030554  
VERSION BC030554.1 GI:20988582  
KEYWORDS HTC.

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 3533)

AUTHORS Strausberg, R.  
TITLE Direct Submission  
JOURNAL Submitted (07-MAY-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.

CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>

Contact: [nisc\\_ungc@ngri.nih.gov](mailto:nisc_ungc@ngri.nih.gov)  
Ahter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,  
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,  
Hansen, N., Ho, S.-L., Karlins, E., Leric, P., Legaas, R., Maduro, O.L.,  
Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J.,  
Pearson, R., Stantrop, S., Thomas, P.J., Touchman, J.W., Tsugeon, C.,  
Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A.,  
Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>

Series: IRAC Plate: 64 Row: k Column: 18  
This clone has the following problem: frame shifted.

Location/Qualifiers  
1..3533  
/organism="Homo sapiens"

## FEATURES

source

1..3533  
/organism="Homo sapiens"

/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5213435"  
/issue\_type="Blood, adult leukocytes"  
/clone\_lib="NIH MGC\_118"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"

## ORIGIN

## Alignment Scores:

Pred. No.: 1.44e-29 Length: 3533  
Score: 307.00 Matches: 58  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-10-031-158-14 (1-58) x BC030554 (1-3533)

Qy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuLysGlnSerSer 20

Db 758 ATGAGATGTTTCCCAAGCCCACTATTTCTTCTTCAATTGCTGAACAAGCTCC 817

Qy 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40

Db 818 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAGATAC 877

Qy 41 TlleglylslslybArgArgAlaThrArpPheTpaPpProArgArgGlyThrPro 58

Db 878 ATTGGAGAGAAAGAGAGACACGATTTCTGGATTCCTGAGAGGGAGACACCA 931

## RESULT 15

LOCUS BU199143 789 bp mRNA linear EST 05-SEP-2002  
DEFINITION DCCBIA08 DCB Homo sapiens cDNA, mRNA sequence.

ACCESSION BU199143  
VERSION BU199143.1 GI:22717723

KEYWORDS EST.  
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 789)  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Xu,X., Gu,J., Liu,F., Qu,J., Zhao,M., Li,Y., Huang,Q., Zhou,J.,  
Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X.,  
Cheng,Z., Xu,S., Gu,W., Tu,Y., Jia,Y., Fu,G., Ren,S., Zhong,M.,  
Lu,G., Cheng,Z. and Han,Z.

TITLE Homo sapiens cDNA DCB clones  
JOURNAL Unpublished (2000)

COMMENT Contact: Zeguang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
201203, P. R. China

Te1: 86-21-50801919(ex.45)  
Fax: 86-21-50801922

Email: hanzg@chgc.sh.cn.

FEATURES  
source Location/Qualifiers

1..789  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/cell\_type="dendritic cells"  
/dev\_stage="mature"  
/lab\_host="BM25.8"  
/clone\_lib="DCB"  
/note="Vector: pTripiEx2; Site\_1: sflta; Site\_2: sfltb"

## ORIGIN

## Alignment Scores:

Pred. No.: 6.11e-30 Length: 789  
Score: 304.00 Matches: 57  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.28% Mismatches: 0

Query Match: 99.02% Indels: 0  
DB: 5 Gaps: 0

US-10-031-158-14 (1-58) x BU199143 (1-789)

Qy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuLysGlnSerSer 20

Db 119 ATGAGATGTTTCCCAAGCCCACTATTTCTTCTTCAATTGCTGAACAAGCTCC 178

Qy 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40

Db 179 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAGATAC 238

Qy 41 TlleglylslslybArgArgAlaThrArpPheTpaPpProArgArgGlyThrPro 58

Db 239 ATTGGAGAGAAAGAGAGACACGATTTCTGGATTCCTGAGAGGGAGACACCA 292

## RESULT 16

LOCUS BE974171 653 bp mRNA linear EST 04-OCT-2000  
DEFINITION 601680522F1 NIH\_MGC\_83 Homo sapiens cDNA clone IMAGE:3950740 5',  
mRNA sequence.

ACCESSION BE974171  
VERSION BE974171.1 GI:10587507

KEYWORDS EST.  
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 653)  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS NIH-MGC http://mgs.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov  
Tissue Procurement: CLONTECH Laboratories, Inc.

CDNA Library Preparation: CLONTECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov

Plate: LUCM817 row: 1 column: 05  
High quality sequence stop: 651.

FEATURES  
source Location/Qualifiers

1..653  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3950740"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NIH MGC\_83"  
/note="Organ: prostate; Vector: pDNR-LIB (Clontech);  
Site\_1: sflit (ggcgctcgcc); Site\_2: sflit  
(ggcattatcgcc); 5' and 3' adaptors were used in cloning  
as follows: 5' adaptor sequence: 5'-CACGGCCATTATGCGC-3'  
and 3' adaptor sequence:  
5'-ATCTAGAGCGGAGGCGGCGGACATG-dt(30)BN-3' (where B = A,  
C, or G and N = A, C, G, or T). Average insert size 1.4  
kb (range 0.5-4.0 kb). 14/15 colonies contained inserts  
by PCR. This library was enriched for full-length clones  
and was constructed by Clontech Laboratories (Palo Alto,  
CA)."

## ORIGIN

## Alignment Scores:

Pred. No.: 1.05e-28 Length: 653  
Score: 294.00 Matches: 58  
Percent Similarity: 98.31% Conservative: 0  
Best Local Similarity: 98.31% Mismatches: 0  
Query Match: 95.77% Indels: 1  
DB: 2 Gaps: 0

US-10-031-158-14 (1-58) x BB974171 (1-653)

Qy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYgInSerSer 20  
 Db 310 ATGCAGATGTTTCCCAAGCCCACTATTTTCTTCCATTCGTAACAAAGCTCC 369  
 Qy 21 ArgArgLeuGlnIshThrPheValPheLeuArgAsnPhseSerLeuMetLeuLeuArgTyr 40  
 Db 370 AGAAGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCGATGTTATTAAGATAC 429

Qy 41 Ile-GlyValysArgArgAlaThrArgPheTyrAspProArgArgGlyThrPro 58  
 Db 430 ATTGGCAAGAAAGAGAGACACATCTCGGATCCGAGGGAACACCA 484

RESULT 17  
 BM919380

LOCUS BM919380 1005 bp mRNA linear EST 12-MAR-2002  
 DEFINITION ACENCCOURT 6715666 NIH\_MGC\_120 Homo sapiens cDNA clone IMAGE:5748632  
 5', mRNA sequence.

ACCESSION BM919380  
 VERSION BM919380.1 GI:19369759  
 KEYWORDS EST.

SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE NIH-MGC http://mgc.nci.nih.gov/  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS NIH-MGC  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabs-remail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.lnl.gov  
 Plate: LHAM1276 row: 1 column: 09  
 High quality sequence stop: 609.

FEATURES  
 source 1..1005  
 location/Qualifiers  
 1..1005  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5748632"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH MGC 120"

/note="Organ: pooled pancreas and spleen; Vector: pCMV-SPORT6; Site 1: NOTI; Site 2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH\_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 1.76e-28 Length: 1005  
 Score: 294.00 Matches: 55  
 Percent Similarity: 98.28% Conservative: 2  
 Best Local Similarity: 94.83% Mismatches: 1  
 Query Match: 95.77% Indels: 0  
 DB: 4 Gaps: 0

US-10-031-158-14 (1-58) x BM919380 (1-1005)

Qy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYgInSerSer 20  
 Db 397 ATGCAGATGTTTCCCAAGCCCACTATTTTCTTCCATTCGTAACAAAGCTCC 456

Qy 21 ArgArgLeuGlnIshThrPheValPheLeuArgAsnPhseSerLeuMetLeuLeuArgTyr 40  
 Db 457 AGAAGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCAAGATATTATTAAGATAC 516

Qy 41 IleGlyValysArgArgAlaThrArgPheTyrAspProArgArgGlyThrPro 58  
 Db 517 ATTGGCAAGAAAGAGAGACACATCTCGGATCCGAGGGAACACCA 570

RESULT 18  
 BC039116

LOCUS BC039116 1757 bp mRNA linear HTC 04-MAR-2003  
 DEFINITION Homo sapiens, similar to T-cell receptor gamma, constant 2, clone  
 IMAGE:4829750, mRNA.

ACCESSION BC039116  
 VERSION BC039116.1 GI:25058606  
 KEYWORDS HTC.

SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE NIH-MGC Project URL: http://mgc.nci.nih.gov  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS NIH-MGC  
 TITLE Direct Submission  
 JOURNAL Submitted (01-NOV-2002) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA

COMMENT Contact: MGC help desk  
 Email: cgabs-remail.nih.gov  
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki  
 Toshiyuki and Piero Carninci (RIKEN)  
 DNA Sequencing by: Baylor College of Medicine Human Genome  
 Sequencing Center  
 Center code: BCM-HQSC  
 Web site: http://www.hgsc.bcm.tmc.edu/cdna/  
 Contact: amg@bcm.tmc.edu  
 Gunaratne, P.H., Garcia, A.M., Lu, X., Huiyk, S.W., Louised, H.,  
 Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,  
 A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov  
 Series: IRAC Plate: 72 Row: f Column: 5  
 This clone has the following problem: retained intron.

FEATURES  
 source 1..1757  
 location/Qualifiers

1..1757  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4829750"  
 /tissue\_type="Testis"  
 /clone\_lib="NIH MGC\_97"  
 /lab\_host="DH10B"  
 /note="Vector: pBluescript"

ORIGIN

Alignment Scores:

Pred. No.: 3.41e-28 Length: 1757  
 Score: 294.00 Matches: 55  
 Percent Similarity: 98.28% Conservative: 2  
 Best Local Similarity: 94.83% Mismatches: 1  
 Query Match: 95.77% Indels: 0  
 DB: 3 Gaps: 0

US-10-031-158-14 (1-58) x BC039116 (1-1757)

Qy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYgInSerSer 20  
 Db 700 ATGCAGATGTTTCCCAAGCCCACTATTTTCTTCCATTCGTAACAAAGCTCC 759

Qy 21 ArgArgLeuGluGlnHisThrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40  
 Db 760 AGAAGGCTGGAACATACCTTGTCTTGTGAGAAATTTTCCAGATATATTAAAGATAC 819  
 Qy 41 IleGlyLeuLeuArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58  
 Db 820 ATTGGCAGAAAAGAGAGCAACGATTCTGGATCCAGAGGGGAGAACCA 873

RESULT 19  
 BF747659/c 434 bp mRNA linear EST 10-JAN-2001  
 DEFINITION QV2-BT0616-051200-521-c07 BT0616 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BF747659  
 VERSION BF747659.1 GI:12074335  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.P., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., deoliveira, P.S., Bucher, P., Jorgensen, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 20202683  
 10737800

JOURNAL MEDLINE PubMed  
 COMMENT Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the PAPSP/LICR Human Cancer Genome project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?l=QV2&t2=QV2-BT0616-051200-521-c07&t3=2000-12-05&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 21  
 High quality sequence stop: 320.  
 Location/Qualifiers  
 1..434  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="BT0616"  
 /note="Organ: breast; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 2,226-28 Length: 434  
 Score: 290.00 Matches: 54  
 Percent Similarity: 98.28% Conservative: 3  
 Best Local Similarity: 93.10% Mismatches: 1  
 Query Match: 94.46% Indels: 0  
 Gaps: 0

US-10-031-158-14 (1-58) x BF747659 (1-434)

Qy 1 MetGlnMetPheProPheProLeuPhePhePheGlnLeuLeuGlnSerSer 20  
 Db 305 ATGCAGATGTTTCCCCCAACCCACTATTTTCTTCTCGATTCGAAACAAACTCC 246

Qy 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40  
 Db 245 AGAAGGCTGGAACATACCTTGTCTTGTGAGAAATTTTCCAGATATATTAAAGATAC 186

Qy 41 IleGlyLeuLeuArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58  
 Db 185 ATTGGCAGAAAAGAGAGCAACGATTCTGGATCCAGAGGGGAGAACCA 132

RESULT 20  
 AM087486/c 497 bp mRNA linear EST 15-OCT-1999  
 LOCUS AM087486  
 DEFINITION XB67C01.X1 Soares NFL T GBC S1 Homo sapiens cDNA clone IMAGE:2581344.3, similar to gb:M13231 T-CELL RECEPTOR GAMMA CHAIN C REGION (HUMAN), mRNA sequence.  
 ACCESSION AM087486  
 VERSION AM087486.1 GI:6043385  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 497)  
 REFERENCE NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
 TITLE Unpublished (1997)  
 JOURNAL COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaaps-rt@mail.nih.gov  
 This clone is available royalty-free through LINT; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 415.  
 Location/Qualifiers  
 1..497  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2581344"  
 /lab\_host="DH10B"  
 /clone\_lib="Soares NFL T GBC S1"  
 /note="Organ: breast; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung KBHL19W, testis NRT, and B-cell NCI-CGAP-GCB1) were mixed, and 88 circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The diver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 2,66-28 Length: 497  
 Score: 290.00 Matches: 54  
 Percent Similarity: 98.28% Conservative: 3  
 Best Local Similarity: 93.10% Mismatches: 1  
 Query Match: 94.46% Indels: 0  
 Gaps: 0

US-10-031-158-14 (1-58) x AM087486 (1-497)

Qy 1 MetGlnMetPheProPheProLeuPhePhePheGlnLeuLeuGlnSerSer 20  
 Db 373 ATGCAGATGTTTCCCCCAACCCACTATTTTCTTCTCGATTCGAAACAAACTCC 314

Qy 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40

```

Db      313 AGAAGCGTGAACATCTTGTCTTGTGAGAAATTTTCCAGATATTTAAGATAC 254
Qy      41 11e61ylyslsAAGATGAlaThrArpHeTtPpPProAGATGAGlyThrPro 58
Db      253 ATTGGCAAGAAAGAAAGAGCCACGATTCGGATTCACAGAGGGGAAACCA 200

RESULT 21
AV714756      681 bp      mRNA      linear      EST 11-OCT-2000
DEFINITION    AV714756 DCB Homo sapiens cDNA clone DCB8C09 5', mRNA sequence.
ACCESSION     AV714756
VERSION       AV714756.1  GI:10796273
KEYWORDS
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE     1 (bases 1 to 681)
AUTHORS       Xu,X., Gu,J., Liu,F., Qu,J., Zhao,M., Li,Y., Huang,Q., Zhou,J.,
Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X.,
Cheng,Z., Xu,S., Gu,W., Tu,Y., Jia,D., Fu,G., Ren,S., Zhong,M.,
Lu,G., Cheng,Z. and Han,Z.
Homo sapiens cDNA DCB clones
Unpublished (2000)
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex. 45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
source
1..681
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DCB8C09"
/cell_type="dendritic cells"
/dev_stage="mature"
/lab_host="BM25.8"
/clone_lib="DCB"
/note="Vector: pTriplex2; Site_1: sf1A; Site_2: sf1B"

ORIGIN
Alignment Scores:
Pred. No.:      3,78e-28      Length:      681
Score:          290.00      Matches:      54
Percent Similarity: 98.28%      Conservative: 3
Best Local Similarity: 93.10%      Mismatches: 1
Query Match:    94.46%      Indels:      0
DB:             1          Gaps:          0

US-10-031-158-14 (1-58) x AV714756 (1-681)

Qy      1  MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuLyGlnSerSer 20
Db      179 ATGCAGATGTTTCCCAAGCCCACTATTTCTTCCATTCGTGAAACAAACTCC 238
Qy      21  ArgArgLeuGlnIsthrPheValPheLeuArgAnpHeSerLeuMetLeuLeuArgTyr 40
Db      239 AGAAGCGTGAACATCTTGTCTTGTGAGAAATTTTCCAGATATTTAAGATAC 298
Qy      41 11e61ylyslsAAGATGAlaThrArpHeTtPpPProAGATGAGlyThrPro 58
Db      299 ATTGGCAAGAAAGAAAGAGCAACGATTCGGATTCACAGAGGGGAAACCA 352

RESULT 22
LOCUS        BC017861      1510 bp      mRNA      linear      HTC 06-DEC-2001
DEFINITION   Homo sapiens, Similar to T cell receptor gamma locus, clone
IMAGE:4687960, mRNA.

```

```

ACCESSION    BC017861
VERSION      BC017861.1  GI:17389678
KEYWORDS
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE     1 (bases 1 to 1510)
AUTHORS       Strausberg,R.
TITLE         Direct Submission
JOURNAL
SUBMITTED (03-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: CLONTECH
cDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site:      http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

FEATURES
source
1..1510
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4687960"
/tissue_type="lung"
/clone_lib="NIH_MGC_77"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"

ORIGIN
Alignment Scores:
Pred. No.:      9,73e-28      Length:      1510
Score:          290.00      Matches:      54
Percent Similarity: 98.28%      Conservative: 3
Best Local Similarity: 93.10%      Mismatches: 1
Query Match:    94.46%      Indels:      0
DB:             3          Gaps:          0

US-10-031-158-14 (1-58) x BC017861 (1-1510)

Qy      1  MetGlnMetPheProSerProLeuPhePhePheLeuGlnLeuLeuLyGlnSerSer 20
Db      479 ATGCAGATGTTTCCCAAGCCCACTATTTCTTCCATTCGTGAAACAAACTCC 538
Qy      21  ArgArgLeuGlnIsthrPheValPheLeuArgAnpHeSerLeuMetLeuLeuArgTyr 40
Db      539 AGAAGCGTGAACATCTTGTCTTGTGAGAAATTTTCCAGATATTTAAGATAC 598
Qy      41 11e61ylyslsAAGATGAlaThrArpHeTtPpPProAGATGAGlyThrPro 58
Db      599 ATTGGCAAGAAAGAAAGAGCAACGATTCGGATTCACAGAGGGGAAACCA 652

RESULT 23
LOCUS        BP964904      1055 bp      mRNA      linear      EST 22-JAN-2001
DEFINITION   602268596F1 NIH_MGC_81 Homo sapiens cDNA clone IMAGE:4357102 5',
mRNA sequence.
ACCESSION     BP964904
VERSION       BP964904.1  GI:12332119
KEYWORDS

```

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
TITLE NIH-MGC http://mgc.nci.nih.gov/  
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
COMMENT Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-remail.nih.gov  
Tissue Procurement: CLONTECH Laboratories, Inc.  
CDNA Library Preparation: CLONTECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.llnl.gov  
Plate: LNCM1217 row: 0 column: 23  
High quality sequence stop: 587.  
Location/Qualifiers  
1..1055  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4357102"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NIH\_MGC\_81"  
/note="Organ: muscle (skeletal); Vector: pDNR-LIB (Clontech); Site 1: Sfil (ggcgccctggcc); Site 2: Sfil (ggccatcggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGCGC-3' and 3' adaptor sequence: 5'-ATCTGAGCGCGGCGGCGGCGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.55 kb (range 1.0-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN  
Alignment Scores:  
Pred. No.: 8,656-28 Length: 1055  
Score: 289.00 Matches: 57  
Percent Similarity: 98.31% Conservative: 1  
Best Local Similarity: 96.61% Mismatches: 0  
Query Match: 94.14% Indels: 1  
DB: 2 Gaps: 0  
US-10-031-158-14 (1-58) x BF675605 (1-1055)

QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20  
DB 67 ATGCAGATGTTTCCCAAGCCCACTATTTCTTCCTTCATTCGTAACAAAGCTCC 126  
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArg-AsnPheSerLeuMetLeuLeuArgTyr 40  
DB 127 AGAAGCGCTGGAACATACCTTGTCTTGTGAAGAACATTTTCCGATGTTATTAAGATA 186  
QY 40 rTlGgLyLyLyLyArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58  
DB 187 CATTGGGAAGAAAGAAAGAACACGATTCGAGATCCCAAGGGAACACCA 241

RESULT 24 523 bp mRNA linear EST 21-DEC-2000  
LOCUS BF675605  
DEFINITION 602083495F1 NIH\_MGC\_83 Homo sapiens CDNA clone IMAGE:4247998 5', mRNA sequence.  
ACCESSION BF675605  
VERSION BF675605.1 GI:11949500  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-remail.nih.gov  
Tissue Procurement: CLONTECH Laboratories, Inc.  
CDNA Library Preparation: CLONTECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.llnl.gov  
Plate: LNCM1065 row: m column: 23  
High quality sequence stop: 510.  
Location/Qualifiers  
1..523  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4247998"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NIH\_MGC\_83"  
/note="Organ: prostate; Vector: pDNR-LIB (Clontech); Site 1: Sfil (ggcgccctggcc); Site 2: Sfil (ggccatcggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGCGC-3' and 3' adaptor sequence: 5'-ATCTGAGCGCGGCGGCGGCGGCGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.4 kb (range 0.5-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN  
Alignment Scores:  
Pred. No.: 5,116-28 Length: 523  
Score: 288.00 Matches: 57  
Percent Similarity: 96.61% Conservative: 0  
Best Local Similarity: 96.61% Mismatches: 1  
Query Match: 93.81% Indels: 1  
DB: 2 Gaps: 0  
US-10-031-158-14 (1-58) x BF675605 (1-523)

QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20  
DB 64 ATGCAGATGTTTCCCAAGCCCACTATTTCTTCCTTCATTCGTAACAAAGCTCC 123  
QY 21 ArgArgLeuGlnHisThrPhe-ValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40  
DB 124 AGAAGCGTGAACATACCTTGTCTTGTGAAGAAATTTTCCGATGTTATTAAGATA 183  
QY 40 rTlGgLyLyLyLyArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58  
DB 184 CATTGGCAAGAAAGAAAGAACACGATTCGAGATCCCAAGGGAACACCA 238

RESULT 25 362 bp mRNA linear EST 07-MAR-2000  
LOCUS A1758378/c  
DEFINITION cy68c12.x1 NCI CGAP Kid11 Homo sapiens CDNA clone IMAGE:2284246 3', similar to gp.M13231 T-CELL RECEPTOR GAMMA CHAIN C REGION (HUMAN);, mRNA sequence.  
ACCESSION A1758378  
VERSION A1758378.1 GI:5152101  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 362)

**AUTHORS** NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
**TITLE** National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
**JOURNAL** Tumor Gene Index  
**COMMENT** Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapdb-remail.nih.gov  
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
 CDNA Library Preparation: M. Bento Soares, Ph.D.  
 DNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: [www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
 Insert Length: 774 Std Error: 0.00  
 Seq primer: -400P from Gldco  
 High quality sequence stop: 286.  
**FEATURES**  
 source  
 1..362  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2284246"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI CGAP Kid1"  
 /note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA from the normalized library NCI-CGAP Kid3 was prepared, and 88 circles were made in vitro. Following NHP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo."

**ORIGIN**  
 Alignment Scores:  
 Pred. No.: 1,13e-27 Length: 362  
 Score: 284.00 Matches: 53  
 Percent Similarity: 96.55% Conservative: 3  
 Best Local Similarity: 91.38% Mismatches: 2  
 Query Match: 92.51% Indels: 0  
 DB: 1 Gaps: 0

US-10-031-158-14 (1-58) x A1758378 (1-362)

Qy 1 MetGlnMetPheProPProSerProLeuPhePhePheLeuGlnLeuLeuLysGlnSerSer 20  
 Db 349 ATGCAGATGTTTCCCCCAGCCCACTATTATCTTCTTCGATGCTGAAACAAACTCC 290

Qy 21 ArgArgLeuGlnIsthrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40  
 Db 289 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCACAGATTTATTAAGATAC 230

Qy 41 IleGlyLysLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58  
 Db 229 ATTGGCAAGAAAGAGAGCAACGATTCGGATCCAGAGAGGGAACACCA 176

**RESULT 26**  
 Bf681238 719 bp mRNA linear EST 21-DEC-2000  
 LOCUS Bf681238 60215502P1 NIH\_MGC\_83 Homo sapiens cDNA clone IMAGE:4296504 5',  
 DEFINITION mRNA sequence.  
 ACCESSION Bf681238  
 VERSION Bf681238.1 GI:11955133  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
**REFERENCE** 1 (bases 1 to 719)  
**AUTHORS** NIH-MGC <http://mgc.ncbi.nlm.nih.gov/>.

**TITLE** National Institutes of Health, Mammalian Gene Collection (MGC)  
**JOURNAL** Unpublished (1999)  
**COMMENT** Contact: Robert Strausberg, Ph.D.  
 Email: cgapdb-remail.nih.gov  
 Tissue Procurement: CLONETECH Laboratories, Inc.  
 CDNA Library Preparation: CLONETECH Laboratories, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: LICM1149 row: c column: 01  
 High quality sequence stop: 673.  
**FEATURES**  
 source  
 1..719  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4296504"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NIH\_MGC\_83"  
 /note="Organ: prostate; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggcgctcgccg); Site 2: SfiI (ggcgctcgccg); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCCGAGGCGCCGACATC-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.4 kb (range 0.5-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

**ORIGIN**  
 Alignment Scores:  
 Pred. No.: 1.61e-26 Length: 719  
 Score: 278.00 Matches: 57  
 Percent Similarity: 98.28% Conservative: 0  
 Best Local Similarity: 98.28% Mismatches: 1  
 Query Match: 90.55% Indels: 1  
 DB: 2 Gaps: 0

US-10-031-158-14 (1-58) x Bf681238 (1-719)

Qy 1 MetGlnMetPheProPProSerProLeuPhePhePheLeuGlnLeuLeuLysGlnSerSer 20  
 Db 68 ATGCAGATGTTTCCCCCAGCCCACTATTATCTTCTTCGATGCTGAAACAAAGCTCC 127

Qy 21 ArgArgLeuGlnIsthrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyr 40  
 Db 128 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCACAGATTTATTAAGATAC 186

Qy 41 IleGlyLysLysArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58  
 Db 187 ATTGGCAAGAAAGAGAGCAACGATTCGGATCCAGAGAGGGAACACCA 240

**RESULT 27**  
 Bf681029 854 bp mRNA linear EST 21-DEC-2000  
 LOCUS Bf681029 602155433P1 NIH\_MGC\_83 Homo sapiens cDNA clone IMAGE:4296104 5',  
 DEFINITION mRNA sequence.  
 ACCESSION Bf681029  
 VERSION Bf681029.1 GI:11954924  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
**REFERENCE** 1 (bases 1 to 854)  
**AUTHORS** NIH-MGC <http://mgc.ncbi.nlm.nih.gov/>.  
**TITLE** National Institutes of Health, Mammalian Gene Collection (MGC)  
**JOURNAL** Unpublished (1999)  
**COMMENT** Contact: Robert Strausberg, Ph.D.

Email: cga@bbs-remail.nih.gov  
Tissue Procurement: CLONETECH Laboratories, Inc.  
cDNA Library Preparation: CLONETECH Laboratories, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <http://image.llnl.gov>  
Plate: L1CM1148 row: b column: 09  
High quality sequence stop: 518.  
Location/Qualifiers

1. 854  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4296104"  
/lab\_host="DH10B (TI phage-resistant)"  
/clone\_id="NIH\_MGC\_83"  
/note="Organ: prostate; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggccgctcgcc); Site 2: SfiI (ggccatgatgccc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCCGAGCGCCGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.4 kb (range 0.5-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

## ORIGIN

Alignment Scores:  
Pred. No.: 2.3e-25 Length: 854  
Score: 270.00 Matches: 53  
Percent Similarity: 93.22% Conservative: 2  
Best Local Similarity: 89.83% Mismatches: 3  
Query Match: 87.95% Indels: 1  
DB: 2 Gaps: 0

US-10-031-158-14 (1-58) x BF681029 (1-854)

QY 1 MetGlnMetPheProSerProLeuPhePhePheGlnLeuLeuGlnSerSer 20  
DB 65 ATGAGATGTTTCCCAAGCCCACTATTTCTTCCTCAATGCTGAACAAAGCTCC 124

QY 21 ArgArgLeuGluHisThr-PheValPheLeuArgAnPheSerLeuMetLeuLeuArgTy 40  
DB 125 AGAAGGCTGGAAACATACCTTTGTTCTTGAAGAAATTTTCCCTGATGTTCAATTAAGAT 184

QY 40 TLeGlyLysLysArgArgAlaThrArgPheThrPheProArgArgGlyThrPro 58  
DB 185 ACATGGCAAGAAAGAAAGACACAGATTCGGATCCAGAGGGGAAACCA 239

RESULT 28  
LOCUS BF679165 629 bp mRNA linear EST 21-DEC-2000  
DEFINITION 602153737P1 NIH\_MGC\_83 Homo sapiens cDNA clone IMAGE:4294761 5',  
mRNA sequence.

ACCESSION BF679165  
KEYWORDS BF679165.1 GI:11953060  
SOURCE EST.  
ORGANISM Homo sapiens (human)

REFERENCE 1 (bases 1 to 629)  
AUTHORS NIH-MGC <http://mcc.nci.nih.gov/>.  
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
COMMENT Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-remail.nih.gov  
Tissue Procurement: CLONETECH Laboratories, Inc.  
cDNA Library Preparation: CLONETECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <http://image.llnl.gov>  
Plate: L1CM1144 row: j column: 10  
High quality sequence stop: 627.  
Location/Qualifiers

1. 629  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4294761"  
/lab\_host="DH10B (TI phage-resistant)"  
/clone\_id="NIH\_MGC\_83"  
/note="Organ: prostate; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggccgctcgcc); Site 2: SfiI (ggccatgatgccc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCCGAGCGCCGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.4 kb (range 0.5-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

## ORIGIN

Alignment Scores:  
Pred. No.: 2.96e-25 Length: 629  
Score: 268.00 Matches: 55  
Percent Similarity: 94.83% Conservative: 0  
Best Local Similarity: 94.83% Mismatches: 3  
Query Match: 87.30% Indels: 1  
DB: 2 Gaps: 0

US-10-031-158-14 (1-58) x BF679165 (1-629)

QY 1 MetGlnMetPheProSerProLeuPhePhePheGlnLeuLeuGlnSerSer 20  
DB 67 ATGAGATGTTTCCCAAGCCCACTATTTCTTCCTCAATGCTGAACAAAGCTCC 126

QY 21 ArgArgLeuGluHisThr-PheValPheLeuArgAnPheSerLeuMetLeuLeuArgTy 40  
DB 127 AGAAGGCTGGAAACATACCTTTGTTCTTGAAGAAATTTTCCCTGATGTTCAATTAAGAT 186

QY 41 TLeGlyLysLysArgArgAlaThrArgPheThrPheProArgArgGlyThrPro 58  
DB 187 AT-GCCAGAAAGAAAGACACAGATTCGGATCCAGAGGGGAAACCA 239

RESULT 29  
LOCUS BF680840 803 bp mRNA linear EST 21-DEC-2000  
DEFINITION 602156088P1 NIH\_MGC\_83 Homo sapiens cDNA clone IMAGE:4297197 5',  
mRNA sequence.

ACCESSION BF680840  
KEYWORDS BF680840.1 GI:11954735  
SOURCE EST.  
ORGANISM Homo sapiens (human)

REFERENCE 1 (bases 1 to 803)  
AUTHORS NIH-MGC <http://mcc.nci.nih.gov/>.  
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
COMMENT Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-remail.nih.gov  
Tissue Procurement: CLONETECH Laboratories, Inc.  
cDNA Library Preparation: CLONETECH Laboratories, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be

Found through the I.M.A.G.E. Consortium/LNL at:  
http://image.jnl.gov  
Plate: LNCMI50 row: 0 column: 22  
High quality sequence stop: 689.  
Location/Qualifiers

## FEATURES

## SOURCE

```
1. 803
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4297197"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_83"
/notes="Organ: prostate; Vector: pDNR-LIB (Clontech);
Site 1: SfiI (ggcgctcgcc); Site 2: SfiI
(ggcattatggcc); 5' and 3' adaptors were used in cloning
as follows: 5' adaptor sequence: 5'-CACGGCCATTATGACC-3'
and 3' adaptor sequence:
5'-ATCTAGAGCGCGGCGCCGACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.4
kb (range 0.5-4.0 kb). 14/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA)."
```

## ORIGIN

## Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
5.38e-25	267.00	803	57	0	1	2	0
Percent Similarity:	96.61%						
Best Local Similarity:	96.61%						
Query Match:	86.97%						

US-10-031-158-14 (1-58) x BF680840 (1-803)

```
QY 1 MetGlnMetPheProPheProLeuPhePhePheLeuGlnLeuLeuGlnSerSer 20
Db 67 ATGCAGATGTTCCCAAGCCCACTATTCTTCTTCCTCAATG-CITGAACAAAGCTCC 125
QY 21 ArgArgLeuGlnIsthrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTy 40
Db 126 AGAAGGCTGGAAACATACCTTGTCTTCTTGAGAAATTTTCCGATGCTTATTAAAGTAC 185
QY 41 Illegl-Lys-Lys-Arg-Arg-Ala-Thr-Arg-Phe-Tyr-Phe-Tyr-Pro-Arg-Gly-Thr-Pro 58
Db 186 ATAGGCAAGAAAGAAAGAGCAACAGATTCTGTGGATCCCAAGAGGGAGACACCA 240
RESULT 30
B1836020 772 bp mRNA linear EST 04-OCT-2001
LOCUS 603084004F1 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5223193 5',
DEFINITION mRNA sequence.
ACCESSION B1836020
VERSION B1836020.1 GI:15947570
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL 1 (bases 1 to 772)
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.jnl.gov
Plate: LNCMI561 row: 9 column: 02
```

High quality sequence stop: 768.  
Location/Qualifiers

## FEATURES

## SOURCE

```
1. 772
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5223193"
/lab_host="DH10B"
/clone_lib="NIH_MGC_120"
/notes="Organ: pooled pancreas and spleen; Vector:
pCW-SpOrt6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of spleen and pancreas from 28 yo
male. Library is oligo-dt primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 1.5 kb, insert size range 1-2.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 025. Note: this is a NIH_MGC library."
```

## ORIGIN

## Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
2.39e-24	262.00	772	54	2	2	2	0
Percent Similarity:	93.33%						
Best Local Similarity:	90.00%						
Query Match:	85.34%						

US-10-031-158-14 (1-58) x B1836020 (1-772)

```
QY 1 MetGlnMetPheProPheProLeuPhePhePheLeuGlnLeuLeuGlnSerSer 20
Db 427 ATGCAGATGTTCCCAAGCCCACTATTCTTCTTCCTCAATGCTGAACAAAGCTCC 486
QY 21 ArgArgLeuGlnIsthrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTy 40
Db 487 AGAAGGCTGGAAACATACCTTGTCTTCTTGAGAAATTTTCCGATATTATTAAGATA 546
QY 40 Illegl-Lys-Lys-Arg-Arg-Ala-Thr-Arg-Phe-Tyr-Phe-Tyr-Pro-Arg-Gly-Thr-Pro 58
Db 547 CATTGGCAACGAAAGAAAGAGCAACAGATTCTGTGGATCCCAAGAGGGAGACACCA 602
RESULT 31
BF673289 808 bp mRNA linear EST 21-DEC-2000
LOCUS 602136139F1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4272944 5',
DEFINITION mRNA sequence.
ACCESSION BF673289
VERSION BF673289.1 GI:11947164
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL 1 (bases 1 to 808)
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
CDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.jnl.gov
Plate: LNCMI087 row: m column: 09
High quality sequence stop: 562.  
Location/Qualifiers
```

## FEATURES

## SOURCE

```
1. 808
/organism="Homo sapiens"
/mol_type="mRNA"
```

/db\_xref="taxon:9606"  
/clone="IMAGE:4272944"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NIH\_MGC\_83"  
/note="Organ: prostate; Vector: pDNR-LIB (Clontech);  
Site\_1: SfiI (ggcgccctggcc); Site\_2: SfiI  
(ggccatcggcc); 5' and 3' adaptors were used in cloning  
as follows: 5' adaptor sequence: 5'-ATTCTAGAGCCGAGCGCCGACATG-dt(30)BN-3' (where B = A,  
C, or G and N = A, C, G, or T). Average insert size 1.4  
kb (range 0.5-4.0 kb). 14/15 colonies contained inserts  
by PCR. This library was enriched for full-length clones  
and was constructed by Clontech Laboratories (Palo Alto,  
CA)."

ORIGIN

Alignment Scores:

Pred. No.:	6.33e-24	Length:	808
Score:	259.00	Matches:	56
Percent Similarity:	93.44%	Conservative:	1
Best Local Similarity:	91.80%	Mismatches:	1
Query Match:	84.36%	Indels:	3
DB:	2	Gaps:	0

US-10-031-158-14 (1-58) x BF673289 (1-808)

Qy 1 MetGlnMetPheProSerProLeuPhePhePheLeuGln-LeuLeuLysGlnSerSe 20  
Db 286 ATGCGAGATGTTCCCGCCAGGCCACTATTTCTTCCTTCAATGCTGCAACAAAGCTCC 345

Qy 20 TArgArgLeuGlnIsthPhe-ValPheLeuArgAsnPheserLeuMetLeuLeuArgT 40  
Db 346 CAGAGGCTGGAAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTAATTAAGAT 405

Qy 40 YrIle-GlyLysLysArgArgAlaThrArgPheTrpAppProArgArgGlyThrPro 58  
Db 406 ACATTGGCAAGAAAGAGAGCAACAGATTCTGGATCCAGAGGGGAAACACCA 462

RESULT 32  
BF674457 857 bp mRNA linear EST 21-DEC-2000  
LOCUS 602137231P1 NIH\_MGC\_83 Homo sapiens cDNA clone IMAGE:4273734 5',  
DEFINITION mRNA sequence.  
ACCESSION BF674457  
VERSION BF674457.1 GI:11948352  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE NIH-MGC http://mgi.nci.nih.gov/  
1 (bases 1 to 857)  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgsab@remail.nih.gov  
Tissue Procurement: CLONETECH Laboratories, Inc.  
cDNA Library Preparation: CLONETECH Laboratories, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L10C1089 row: n column: 07  
High quality sequence stop: 638.  
Location/Qualifiers  
1..857  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4273734"  
/lab\_host="DH10B (T1 phage-resistant)"

FEATURES  
source

/clone\_lib="NIH\_MGC\_83"  
/note="Organ: prostate; Vector: pDNR-LIB (Clontech);  
Site\_1: SfiI (ggcgccctggcc); Site\_2: SfiI  
(ggccatcggcc); 5' and 3' adaptors were used in cloning  
as follows: 5' adaptor sequence: 5'-ATTCTAGAGCCGAGCGCCGACATG-dt(30)BN-3' (where B = A,  
C, or G and N = A, C, G, or T). Average insert size 1.4  
kb (range 0.5-4.0 kb). 14/15 colonies contained inserts  
by PCR. This library was enriched for full-length clones  
and was constructed by Clontech Laboratories (Palo Alto,  
CA)."

ORIGIN

Alignment Scores:

Pred. No.:	4.29e-23	Length:	857
Score:	253.00	Matches:	56
Percent Similarity:	96.55%	Conservative:	0
Best Local Similarity:	96.55%	Mismatches:	2
Query Match:	82.41%	Indels:	2
DB:	2	Gaps:	0

US-10-031-158-14 (1-58) x BF674457 (1-857)

Qy 1 MetGlnMetPheProSerProLeuPhePhePheLeuGlnLeuLeuLysGlnSerSe 20  
Db 63 ATGCGAGATGTTCCCGCCAGGCCACTATTTCTTCCTTCAATG-CTGAAACAAAGCTCC 121

Qy 21 TArgArgLeuGlnIsthPhe-ValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40  
Db 122 AGAAGGCTGGAAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTAATTAAGATC 181

Qy 41 IlegLysLysArgArgAlaThrArgPheTrpAppProArgArgGlyThrPro 58  
Db 182 AT-GCGAAGAAAGAGAGCAACAGATTCTGGATCCAGAGGGGAAACACCA 234

RESULT 33  
BF679284 849 bp mRNA linear EST 21-DEC-2000  
LOCUS 602153477P1 NIH\_MGC\_83 Homo sapiens cDNA clone IMAGE:4294386 5',  
DEFINITION mRNA sequence.  
ACCESSION BF679284  
VERSION BF679284.1 GI:11953179  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE NIH-MGC http://mgi.nci.nih.gov/  
1 (bases 1 to 849)  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgsab@remail.nih.gov  
Tissue Procurement: CLONETECH Laboratories, Inc.  
cDNA Library Preparation: CLONETECH Laboratories, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L10C1143 row: j column: 19  
High quality sequence stop: 587.  
Location/Qualifiers  
1..849  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4294386"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NIH\_MGC\_83"  
/note="Organ: prostate; Vector: pDNR-LIB (Clontech);  
Site\_1: SfiI (ggcgccctggcc); Site\_2: SfiI

FEATURES  
source

```

/db.xref="taxon:9544"
/clone="IBUW:23081"
/sex="male"
/tissue_type="blood"
/cell_type="PBMC"
/dev_stage="adult"
/lab_host="Electromax DH10B"
/clone_id="Katzu.FMPB2"
[note="Vector: pDONR 221, Site 1: Berg I, Site 2: Berg I, Site 3: Berg I"]

```

### Alignment Scores:

Alignment Scores:	
Pred. No.:	2,38e-22
Score:	248.00
Percent Similarity:	89.66%
Best Local Similarity:	84.48%
Length:	995
Matches:	49
Conservative:	3
Mismatches:	6

DB:	7	Gap
-----	---	-----

US-10-031-158-14 (1-58) X C0647648 (1-995)

1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLysGlnSerSer 20

Db 468 ATGCAGATGTTTCACCCCAAGCCCACTATTCTTCTCCATTGCTGAACAACCTCC 52

21 ArgArgLeuGluHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40

Db 528 ACAAGGCTGGACATACCTTGTCTTCTTGAGAAATTTTCCGAGATGTTATTGAGAT

41 IleglylsArgAlaThrArgPheTrpAspProArgGlyThrPro 58

Db 588 ATTGGCAGAAAGAACGACACACAGGTTCTGAATCCCAAGAGGGGAAACACCA 641

RESULT 35

LOCUS	CB553506	663 bp	mRNA	linear	EST 01-JUN-2003
-------	----------	--------	------	--------	-----------------

ACCESSION CB553506

**KEYWORDS:** EST.

ORGANISM *Macaca mulatta*

Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;

REFERENCE 1 (bases 1 to 663)

Holzman, T.

JOURNAL  
Unpublished (2002)

Katze Lab

Box 358070, Seattle, WA 98195-8070, USA

Fax: 206 732 6055

Similar to GenBank entry HSTRGC10 X06775 Human aberrant mRNA from

Plate: MMSp0044 row: C column: 12.

Source 1. .663

```
/mol_type="mrna"
```

```
/sex="male"
```

```

/cell_type="monocyte"
/dev_stage="adult"
/clone_lib="MMP"
/note="Organ: spleen"

```

Alignment Scores:

Pred. No.:	Length:	Score:	Matches:	Conservative:	Mismatches:	Best Local Similarity:	Indels:	Query Match:
3.69e-22	663	245.00	48	89.66%	82.76%	79.80%	0	6

US-10-031-158-14 (1-58) x CB553506 (1-663)

Qy 1 MetGlnMetPheProPsoSerProLeuPhePheLeuGlnLeuLeuLyGlnSerSer 20  
 Db 401 ATGCGAGATGTTCCACCCAGCCCACTATATTTCTTCTTCATTGCTGAACAAACCTCC 460  
 Qy 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTy 40  
 Db 461 ACAAGCGTGAACATACCTTTGTCTTCTTGAGAAATTTTCCAGATGTTATTTAGATAC 520  
 Qy 41 TlGlyLyLyBaArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58  
 Db 521 ATTGGCAGAAAGAAAGAACAGCAAGATTCTGAATATCCCAAGAGGGAACACCA 574

RESULT 36  
 BF678129 870 bp mRNA linear EST 21-DEC-2000  
 LOCUS 602085181P1 NIH\_MGC\_83 Homo sapiens cDNA clone IMAGE:4249245 5',  
 DEFINITION mRNA sequence.  
 BF678129  
 VERSION BF678129.1 GI:11952024  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 870)  
 NIH-MGC http://mgs.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: rgs@bhs-remail.nih.gov  
 Tissue Procurement: CLONETECH Laboratories, Inc.  
 cDNA Library Preparation: CLONETECH Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNLN at:  
 http://image.llnl.gov  
 Plate: LNCM1069 row: a column: 22  
 High quality sequence etop: 557.

# FEATURES

Location/Qualifiers

1..870  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4249245"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NIH MGC 83"  
 /note="Organ: prostate; Vector: pDNR-LIB (Clontech);  
 Site 1: Sfil (ggcgctcgcc); Site 2: Sfil  
 (ggcgctcgcc); 5' and 3' adaptors were used in cloning  
 as follows: 5' adaptor sequence: 5'-CACGCGCTTATGCGC-3'  
 and 3' adaptor sequence:  
 5'-ATTCTAGAGCGCGAGCGCGACATG-dt(30)BN-3' (where B = A,  
 C, or G and N = A, C, G, or T). Average inserted size 1.4  
 kb (range 0.5-4.0 kb). 14/15 colonies contained inserts  
 by PCR. This library was enriched for full-length clones  
 and was constructed by Clontech Laboratories (Palo Alto,  
 CA)."

# ORIGIN

Alignment Scores:

1.28e-19 Length: 870

Score: 227.00 Matches: 56  
 Percent Similarity: 93.33% Conservative: 0  
 Best Local Similarity: 93.33% Mismatches: 2  
 Query Match: 73.94% Indels: 4  
 DB: 2 Gaps: 0

US-10-031-158-14 (1-58) x BF678129 (1-870)

Qy 1 MetGlnMetPheProPsoSerProLeuPhePheLeuGlnLeuLeuLyGlnSerSer 20  
 Db 63 ATGCGAGATGTTCCACCCAGCCCACTATATTTCTTCTTCATTGCTGAACAAACCTCC 121  
 Qy 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTy 40  
 Db 122 AGAAGCGTGAACATACCTTTGTCTTCTTGAGAAATTTTCCAGATGTTATTTAGATAC 181  
 Qy 40 TlGlyLyLyBaArgArgAlaThrArgPheTrpAspProArgArgGlyThrPro 58  
 Db 182 CAT-GGGAAGAAAGAAAGAACAGCAAGATTCTGGGATCCCAAGAGGGAACACCA 236

RESULT 37  
 CB551454 637 bp mRNA linear EST 01-JUN-2003  
 LOCUS MMSP0013 E12 MMSP Macaca mulatta cDNA, mRNA sequence.  
 DEFINITION CB551454  
 ACCESSION CB551454.1 GI:31300649  
 VERSION CB551454.1 GI:31300649  
 KEYWORDS EST.  
 SOURCE Macaca mulatta (rhesus monkey)  
 ORGANISM Macaca mulatta  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 Cercopithecinae; Macaca.

REFERENCE 1 (bases 1 to 637)  
 Katze, M.G., Bungarner, R., Korth, M., Feldman, R., Amjadi, M. and  
 Holzman, T.  
 TITL Expresed sequence tags from Rhesus macaque spleen  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Holzman T  
 Katze Lab  
 University of Washington  
 Box 358070, Seattle, WA 98195-8070, USA  
 Tel: 206 732 6156  
 Fax: 206 732 6055  
 Email: ted@locke.hb.washington.edu  
 Similar to Genbank entry S60779 S60779 Homo sapiens nonfunctional  
 T-cell receptor gamma chain variable region mRNA, partial sequence.  
 8/2000  
 Plate: MMSP0013 row: B column: 12.

# FEATURES

Location/Qualifiers

1..637  
 /organism="Macaca mulatta"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9544"  
 /sex="male"  
 /cell\_type="monocytes"  
 /dev\_stage="adult"  
 /clone\_lib="MMSP"  
 /note="Organ: spleen"

# ORIGIN

Alignment Scores:

Pred. No.:	Length:	Score:	Matches:	Conservative:	Mismatches:	Best Local Similarity:	Indels:	Query Match:
1.21e-19	637	226.00	48	86.44%	81.36%	73.62%	0	6

US-10-031-158-14 (1-58) x CB551454 (1-637)

Qy 1 MetGlnMetPheProPsoSerProLeuPhePheLeuGlnLeuLeuLyGlnSerSer 20  
 Db 454 ATGCGAGATGTTCCACCCAGCCCACTATATTTCTTCTTCATTGCTGAACAAACCTCC 513

Qy 21 ArgArgLeuGluHisThrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40  
 Db 514 ATAGGCTGGACATACCTTTCTCTTGGAAATTTTCCAGATGTTATTATGAGATAC 573  
 Qy 41 -11eg1LyLyLySArGAlaThrArGPhetPAsPProArGArgLyThrPro 58  
 Db 574 CATTGGCAGAAAGAAAGACAGACAGAGTTCTGAATCCAGAGGGGGAACACCA 628

RESULT 38  
 BF677916 670 bp mRNA linear EST 21-DEC-2000  
 LOCUS 602084736F1 NIH\_MGC\_83 Homo sapiens cDNA clone IMAGE:4248993 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BF677916 GI:11951811  
 VERSION BF677916.1  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 670)  
 NIH-MGC http://mgi.nci.nih.gov/.  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished (1999)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: cgabs-remail.nih.gov  
 Tissue Procurement: CLONTECH Laboratories, Inc.  
 CDNA Library Preparation: CLONTECH Laboratories, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNLN at:  
 http://image.llnl.gov  
 Plate: L1CM1068 row: 9 column: 10  
 High quality sequence stop: 668.

## FEATURES

location/Qualifiers  
 1..670  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4248993"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NIH\_MGC\_83"  
 /note="Organ: prostate; Vector: pDNR-LIB (Clontech);  
 Site 1: Sf1 (ggcgctcgcc); Site 2: Sf1  
 (ggcattatggcc); 5' and 3' adaptors were used in cloning  
 as follows: 5' adaptor sequence: 5'-CACGCCATTATGCGC-3'  
 and 3' adaptor sequence:  
 5'-ATTCTAAGGCCGAGCGCGCGACATG-dt(30)BN-3' (where B = A,  
 C, or G and N = A, C, G, or T). Average insert size 1.4  
 kb (range 0.5-4.0 kb). 14/15 colonies contained inserts  
 by PCR. This library was enriched for full-length clones  
 and was constructed by Clontech Laboratories (Palo Alto,  
 CA)."

## ORIGIN

Alignment Scores:  
 Pred. No.: 2,37e-19 Length: 670  
 Score: 224.00 Matches: 55  
 Percent Similarity: 94.83% Conservative: 0  
 Best Local Similarity: 94.83% Mismatches: 3  
 Query Match: 72.96% Indels: 3  
 DB: 2 Gaps: 0

US-10-031-158-14 (1-58) x BF677916 (1-670)

Qy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20  
 Db 65 ATGAGATGTTTCCCAAGCCCACTATTTCTTCCTTCATG-CTGAAACAAAGCTCC 123  
 Qy 21 ArgArgLeuGluHisThrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40  
 Db 40 yrl1eg1LyLySArGAlaThr-ArgPhetPAsPProArGArgLyThrPro 58

Db 124 AGAAGCTGGACATACCTT-GTCTTCTGGAAATTTTCCCTGATTTATTAGATAC 182  
 Qy 41 11eg1LyLyLySArGAlaThrArGPhetPAsPProArGArgLyThrPro 58  
 Db 183 AT-GGCAGAAAGAAAGACAGACAGATTTGGATTCAGAGGGGGAACACCA 235

RESULT 39  
 BG535408 827 bp mRNA linear EST 03-APR-2001  
 LOCUS 602563069F1 NIH\_MGC\_77 Homo sapiens cDNA clone IMAGE:4687960 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG535408  
 VERSION BG535408.1 GI:13526953  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 827)  
 NIH-MGC http://mgi.nci.nih.gov/.  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished (1999)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: cgabs-remail.nih.gov  
 Tissue Procurement: CLONTECH Laboratories, Inc.  
 CDNA Library Preparation: CLONTECH Laboratories, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNLN at:  
 http://image.llnl.gov  
 Plate: L1CM1501 row: 1 column: 17  
 High quality sequence stop: 609.

## FEATURES

location/Qualifiers  
 1..827  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4687960"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NIH\_MGC\_77"  
 /note="Organ: lung; Vector: pDNR-LIB (Clontech); Site 1:  
 Sf1 (ggcgctcgcc); Site 2: Sf1 (ggcattatggcc); 5' and  
 3' adaptors were used in cloning as follows: 5' adaptor  
 sequence: 5'-CACGCCATTATGCGC-3' and 3' adaptor sequence:  
 5'-ATTCTAAGGCCGAGCGCGCGACATG-dt(30)BN-3' (where B = A,  
 C, or G and N = A, C, G, or T). Average insert size 1.9  
 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts  
 by PCR. This library was enriched for full-length clones  
 and was constructed by Clontech Laboratories (Palo Alto,  
 CA). Note: this is a NIH\_MGC Library."

## ORIGIN

Alignment Scores:  
 Pred. No.: 7.64e-19 Length: 827  
 Score: 221.00 Matches: 50  
 Percent Similarity: 88.71% Conservative: 5  
 Best Local Similarity: 80.65% Mismatches: 3  
 Query Match: 71.99% Indels: 4  
 DB: 4 Gaps: 0

US-10-031-158-14 (1-58) x BG535408 (1-827)

Qy 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuGlnSerSer 20  
 Db 489 ATGAGATGTTTCCCAAGCCCACTATTTCTTCCTTCGATTCGAAACAAACTCC 548  
 Qy 21 ArgArgLeuGluHisThrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40  
 Db 549 AGAAGCTGGACATACCTTTGCTTCTTGAAGAAATATTCACAGATATTTAAGCAT 608  
 Qy 40 yrl1eg1LyLySArGAlaThr-ArgPhetPAsPProArGArgLyThrPro 58  
 Db 40 yrl1eg1LyLySArGAlaThr-ArgPhetPAsPProArGArgLyThrPro 58



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comphen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 8, 2004, 07:51:31 ; Search time 56 Seconds

(without alignments)  
736.174 Million cell updates/sec

Title: US-10-031-158-14

Perfect score: 307  
Sequence: 1 MOMPSPFLFQLKQSS.....RYGKKRRATRFMDPRGTP 58

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 824507 seqs, 353594441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame\_p2n.model -DEV=x1h  
-Q=/cgn2\_1/USPTO.epool/US10031158/runat\_06122004\_082700\_15922/app\_query.fasta\_1.199  
-DB=Issued\_Patents\_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blousum62 -TRANS=human40.cdi  
-LIST=45 -DOCALL=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10031158 @CGN1\_1\_54@runat\_06122004\_082700\_15922 -NCP=6 -ICPU=3  
-NO\_MMAP -LARGEOQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued\_Patents\_NA:\*  
1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/prodata/1/ina/5A\_COMB.seq:\*  
4: /cgn2\_6/prodata/1/ina/5B\_COMB.seq:\*  
5: /cgn2\_6/prodata/1/ina/BCTUS\_COMB.seq:\*  
6: /cgn2\_6/prodata/1/ina/backfillseq1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	307	100.0	426	4	US-09-513-999C-679
2	307	100.0	825	1	US-08-256-964A-18
3	279	90.9	1080	6	5260223-3
4	251	81.8	460	4	US-09-513-999C-8197
5	65.5	21.3	1789	4	US-09-142-108C-1
6	64	20.8	503	4	US-09-250-609-85
7	64	20.8	503	4	US-09-250-611-85
8	63.5	20.7	74962	4	US-09-685-853A-3
9	63	20.5	195	4	US-09-248-796A-10951
10	62.5	20.4	234	4	US-09-248-796A-11408
11	62.5	20.4	4070	3	US-09-302-812-1
12	62.5	20.4	4070	3	US-09-511-477-1

13	62.5	20.4	4070	3	US-09-511-507-1	Sequence 1, Appl1
14	61.5	20.0	2755	3	US-09-025-769B-274	Sequence 274, App
15	61.5	20.0	2755	4	US-09-490-070A-274	Sequence 274, App
16	61.5	20.0	2755	4	US-09-490-153-274	Sequence 274, App
17	61	19.9	339	4	US-09-250-609-79	Sequence 79, Appl
18	61	19.9	339	4	US-09-250-611-79	Sequence 79, Appl
19	61	19.9	399	4	US-09-250-609-91	Sequence 91, Appl
20	61	19.9	399	4	US-09-250-611-91	Sequence 91, Appl
21	61	19.9	420	4	US-09-250-609-89	Sequence 89, Appl
22	61	19.9	420	4	US-09-250-611-89	Sequence 89, Appl
23	61	19.9	424	4	US-09-250-609-77	Sequence 77, Appl
24	61	19.9	424	4	US-09-250-611-77	Sequence 77, Appl
25	61	19.9	471	4	US-09-250-609-90	Sequence 90, Appl
26	61	19.9	471	4	US-09-250-611-90	Sequence 90, Appl
27	61	19.9	540	4	US-09-250-609-95	Sequence 95, Appl
28	61	19.9	540	4	US-09-250-611-95	Sequence 95, Appl
29	61	19.9	836	4	US-09-495-050A-103	Sequence 103, App
30	61	19.9	2136	4	US-09-587-184-1	Sequence 1, Appl1
31	61	19.9	3798	4	US-09-688-188B-9	Sequence 9, Appl1
32	61	19.9	3798	4	US-09-291-417D-9	Sequence 9, Appl1
33	61	19.9	4266	3	US-09-651-011A-3	Sequence 3, Appl1
34	61	19.9	7295	2	US-08-487-826B-15	Sequence 15, Appl1
35	60.5	19.7	219	4	US-09-328-352-3224	Sequence 3224, Ap
36	60.5	19.7	504	4	US-09-134-000C-3042	Sequence 3042, Ap
37	60.5	19.7	1463	3	US-08-943-731-7	Sequence 7, Appl1
38	60.5	19.7	18609	4	US-08-943-681A-192	Sequence 1, Appl1
39	60	19.5	2715	4	US-09-534-638-1	Sequence 1, Appl1
40	60	19.5	30350	3	US-10-118-328-3	Sequence 3, Appl1
41	59.5	19.4	357	4	US-09-513-999C-15620	Sequence 15620, A
42	59.5	19.4	1144	4	US-09-149-476-128	Sequence 128, App
43	59.5	19.4	1482	4	US-09-023-655-191	Sequence 191, App
44	59.5	19.4	2095	4	US-09-608-917A-1	Sequence 1, Appl1
45	59.5	19.4				

#### ALIGNMENTS

RESULT 1  
US-09-513-999C-679  
; Sequence 679, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Mline Edwards, J.B.  
; APPLICANT: Duciery, A.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59, US2, REG  
; CURRENT APPLICATION NUMBER: US/09/513, 999C  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 679  
; LENGTH: 426  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 267..425  
US-09-513-999C-679  
Alignment Scores:  
Pred. No.: 1.64e-36  
Score: 307.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 4  
Gaps: 0  
US-10-031-158-14 (1-58) x US-09-513-999C-679 (1-426)

QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYsgInSerSer 20  
 DB 94 ATGCAGATGTTTCCCAAGCCCACTATTCTTCTTCAATGCTGAAACAAAGCTCC 153  
 QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40  
 DB 154 AGAAGGCTGGAACATACCTTTGCTTCTTGAGAAATTTTCCCTGATGTTATTAAAGATAC 213  
 QY 41 IlegLYLyeLYsArgArgAlaThrArgPheTyrAspProArgArgGlyThrPro 58  
 DB 214 ATTGGCAAGAAAGAAAGAGCAACACGATTCTGGGATCCCAAGAGGGGAACACCA 267

## RESULT 2

US-08-256-964A-18  
 ; Sequence 18, Application US/08256964A  
 ; Patent No. 5723309

## GENERAL INFORMATION:

APPLICANT: BONNEVILLE, Marc  
 TITLE OF INVENTION: PRODUCTION OF SUBUNITS OF SOLUBLE T  
 TITLE OF INVENTION: RECEPTORS BY CO-TRANSFECTION, USES OF THE PRODUCTS THUS  
 TITLE OF INVENTION: OBTAINED  
 NUMBER OF SEQUENCES: 19  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Young & Thompson  
 STREET: 745 South 23rd Street  
 CITY: Arlington  
 STATE: VA  
 COUNTRY: USA  
 ZIP: 22202  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/256,964A  
 FILING DATE: 14-SEP-1994  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: FR 92 14203  
 FILING DATE: 25-NOV-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: PATCH, Andrew J.  
 REGISTRATION NUMBER: Reg. No. 5723309 32,925  
 REFERENCE/DOCKET NUMBER: BE 94/449  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 703/521-2297  
 TELEFAX: 703/685-0573  
 TELEX: 248425

## INFORMATION FOR SEQ ID NO: 18:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 825 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 FEATURE:

## NAME/KEY: CDS

LOCATION: 1..825  
 PUBLICATION INFORMATION:  
 DOCUMENT NUMBER: WO 94/12648  
 FILING DATE: 25-NOV-1993  
 PUBLICATION DATE: 09-JUN-1994  
 US-08-256-964A-18

## Alignment Scores:

Pred. No.: 4.18e-36  
 Score: 307.00  
 Percent Similarity: 100.00%  
 Best Local Similarity: 100.00%  
 Query Match: 100.00%  
 DB: 1

## Length: 825

Matches: 58  
 Conservative: 0  
 Mismatches: 0  
 Indels: 0  
 Gaps: 0

US-10-031-158-14 (1-58) x US-08-256-964A-18 (1-825)  
 QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYsgInSerSer 20  
 DB 440 ATGCAGATGTTTCCCAAGCCCACTATTCTTCTTCAATGCTGAAACAAAGCTCC 499  
 QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40  
 DB 500 AGAAGGCTGGAACATACCTTTGCTTCTTGAGAAATTTTCCCTGATGTTATTAAAGATAC 559  
 QY 41 IlegLYLyeLYsArgArgAlaThrArgPheTyrAspProArgArgGlyThrPro 58  
 DB 560 ATTGGCAAGAAAGAAAGAGCAACACGATTCTGGGATCCCAAGAGGGGAACACCA 613

## RESULT 3

5260223-3  
 ; Patent No. 5260223

APPLICANT: BRENNER, MICHAEL B.; STROMINGER, JACK L.; SEIDMAN,  
 JOHN G., JR., STEPHEN H., KRANGEL, MICHAEL S.  
 TITLE OF INVENTION: METHODS FOR DETECTION OF HUMAN GAMMA  
 T CELL RECEPTOR

## NUMBER OF SEQUENCES: 4

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/187,698  
 FILING DATE: 29-APR-1998  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 115,256  
 FILING DATE: 29-OCT-1987  
 APPLICATION NUMBER: 16,252  
 FILING DATE: 19-FEB-1987  
 APPLICATION NUMBER: 882,100  
 FILING DATE: 03-JUL-1986  
 SEQ ID NO: 3  
 LENGTH: 1080

5260223-3

## Alignment Scores:

Pred. No.: 8.8e-32  
 Score: 279.00  
 Percent Similarity: 96.55%  
 Best Local Similarity: 91.38%  
 Query Match: 90.88%  
 DB: 6  
 Length: 1080  
 Matches: 53  
 Conservative: 3  
 Mismatches: 2  
 Indels: 0  
 Gaps: 0

US-10-031-158-14 (1-58) x 5260223-3 (1-1080)

QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYsgInSerSer 20  
 DB 452 ATGCAGATGTTTCCCAAGCCCACTATTCTTCTTCAATGCTGAAACAAAGCTCC 511  
 QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40  
 DB 512 AGAAGGCTGGAACATACCTTTGCTTCTTGAGAAATTTTCCCAATATTATTAAAGATAC 571  
 QY 41 IlegLYLyeLYsArgArgAlaThrArgPheTyrAspProArgArgGlyThrPro 58  
 DB 572 ATTGGCAAGAAAGAAAGAGCAACACGATTCTGGGATCCCAAGAGGGGAACACCA 625

## RESULT 4

US-09-513-999C-8197  
 ; Sequence 8197, Application US/09513999C  
 ; Patent No. 6783961

## GENERAL INFORMATION:

APPLICANT: Dumas, Milne Edwards, J.B.  
 APPLICANT: Duardet, A.  
 APPLICANT: Giordano, J.Y.  
 TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
 FILE REFERENCE: 59,US2,REG  
 CURRENT APPLICATION NUMBER: US/09/513,999C  
 CURRENT FILING DATE: 2000-02-24  
 PRIOR APPLICATION NUMBER: US 60/122,487  
 PRIOR FILING DATE: 1999-02-26

```

; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patenc.pm
; SEQ ID NO 8197
; LENGTH: 460
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 327
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 345
; OTHER INFORMATION: w=a or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 421
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 422
; OTHER INFORMATION: n=a, g, c or t
; US-09-513-999C-8197

Alignment Scores:
Pred. No.: 3,78e-28      Length: 460
Score: 251.00           Matches: 56
Percent Similarity: 96.53%      Conservative: 0
Best Local Similarity: 96.55%      Mismatches: 2
Query Match: 81.76%             Indels: 2
DB: 4                     Gaps: 0

US-10-031-158-14 (1-58) x US-09-513-999C-8197 (1-460)

QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYsgInserSer 20
   |||||
Db 200 ATGCAGATGTTTCCCAAGCCACATATTTCTTCTTCAATGCTGAACAAAGCTC- 258

QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPhSerLeuMetLeuLeuArgTy 40
   |||||
Db 259 ACAAAGGCTGGAACATCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAGATAC 318

QY 41 IleglyLysLysArgArgAlaThrArgPheTTPAspProArgArgGlyThrPro 58
   |||||
Db 319 ATTGGGAARAAAGAGAGCAACAGTC-TGGGATCCAGAGGAGGAAACCA 371

RESULT 5
US-09-142-108C-1/c
; Sequence 1, Application US/09142108C
; Patent No. 6774285
; GENERAL INFORMATION:
; APPLICANT: Bruggiera, Filippo
; APPLICANT: Holton, Timothy A.
; APPLICANT: Michael, Michael Z.
; TITLE OF INVENTION: GENETIC SEQUENCES ENCODING FLAVONOID PATHWAY ENZYMES
; FILE REFERENCE: 11658
; CURRENT APPLICATION NUMBER: US/09/142,108C
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: PN8386
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Patencin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1789
; TYPE: DNA
; ORGANISM: Petunia x hybrida
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (50) .. (1588)
; US-09-142-108C-1

Alignment Scores:
```

```

Pred. No.: 9.03      Length: 1789
Score: 65.50         Matches: 18
Percent Similarity: 63.04%      Conservative: 11
Best Local Similarity: 39.13%      Mismatches: 14
Query Match: 21.34%             Indels: 4
DB: 4                 Gaps: 1

US-10-031-158-14 (1-58) x US-09-142-108C-1 (1-1789)

QY 4 PheProProSerProLeuPhePheLeuGlnLeuLeuYsgInserSerArgArgLeu 23
   |||||
Db 912 TTCCCTTCATCAT-ATCCGATCATCTTTTAAAGAGATCAAAAGTCAACAAATTC 854

QY 24 GlnHisThrPheValPheLeuArgAsnPhSerLeuMetLeuLeuArgTyIleglyLys 43
   |||||
Db 853 TTTCA------TTCTCCAAAAATTTTACCTTATGCTCTTCAAGTATATCACTCA 803

QY 44 LysArgArgAlaThrArg 49
   |||||
Db 802 GAAGCGCTCGAAACGCCG 785

RESULT 6
US-09-250-609-85
; Sequence 85, Application US/09250609A
; Patent No. 6458943
; GENERAL INFORMATION:
; APPLICANT: Byrne, Jennifer A.
; FILE REFERENCE: 1383.0210002
; CURRENT APPLICATION NUMBER: US/09/250,609A
; CURRENT FILING DATE: 1999-02-17
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: Patencin Ver. 2.0
; SEQ ID NO 85
; LENGTH: 503
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-250-609-85

Alignment Scores:
Pred. No.: 2.5      Length: 503
Score: 64.00       Matches: 13
Percent Similarity: 59.26%      Conservative: 3
Best Local Similarity: 48.15%      Mismatches: 11
Query Match: 20.85%             Indels: 0
DB: 4                 Gaps: 0

US-10-031-158-14 (1-58) x US-09-250-609-85 (1-503)

QY 31 ArgAsnPhSerLeuMetLeuLeuArgTyIleglyLysLysArgArgAlaThrArgPhe 50
   |||||
Db 246 AGGAGCTTCGGGCTAGCTGTCTAAGTGGAAGAAATTTGTCACTCTGCCCGGCTGC 305

QY 51 TTPAspProArgArgGlyThr 57
   |||||
Db 306 TCGACGCCAAGAGAGGCACT 326

RESULT 7
US-09-250-611-85
; Sequence 85, Application US/09250611
; Patent No. 6528283
; GENERAL INFORMATION:
; APPLICANT: Byrne, Jennifer A.
; APPLICANT: Bassett, Paul
; TITLE OF INVENTION: Members of the D52 Gene Family
; FILE REFERENCE: 1383.0210001
; CURRENT APPLICATION NUMBER: US/09/250,611
; CURRENT FILING DATE: 1999-02-17
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: Patencin Ver. 2.0
; SEQ ID NO 85
; LENGTH: 503
; TYPE: DNA
```



```

Db      91 TTTTITTTTTAGATTCATTCTTCTTTGTTGTTATTTAGAAAAAACCATA 150
        |||          ::|||   ||::|||   ||::|||   ||::|||
Qy      48 ThrArgPheTp 51
        |||          |||::|||
Db      151 CCAAGATTTTT 162

RESULT 11
US-09-302-812-1
; Sequence 1, Application US/09302812B
; Patent No. 6333148
; GENERAL INFORMATION:
; APPLICANT: JACOBSON, Myron K.
; APPLICANT: JACOBSON, Elaine L.
; APPLICANT: AM, Jean-Christophe
; APPLICANT: LIN, Winston
; TITLE OF INVENTION: GENES ENCODING SEVERAL POLY(ADP-RIBOSE) GLYCOPHYROLASE (PARG) ENZYME COMPOSITIONS AND ANTIBODIES IMMUNOREACTIVE
; TITLE OF INVENTION: THE PROTEINS AND FRAGMENTS THEREOF, AND ANTIBODIES IMMUNOREACTIVE
; FILE REFERENCE: NIND 201
; CURRENT APPLICATION NUMBER: US/09/302,812B
; EARLIER FILING DATE: 1999-04-30
; EARLIER FILING DATE: 1998-05-01
; NUMBER OF SEQ ID NOS: 38
; SEQ ID NO 1
; LENGTH: 4070
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
US-09-302-812-1

Alignment Scores:
Pred. No.:           80.7           Length:         4070
Score:              62.50          Matches:         17
Percent Similarity: 53.85%         Conservative:    4
Best Local Similarity: 43.59%       Mismatches:     17
Query Match:        20.36%         Indels:         1
DB:                  3             Gaps:            1
US-10-031-158-14 (1-58) x US-09-302-812-1 (1-4070)

Qy      1 MetGlnMetPheProPheSerProLeuPhePheLeuGlnLeuLysGlnSerSer 20
        |||||          |||||          |||||          |||||
Db      3393 ATGAGAGTCCTCCCGCCCATGCAGCTGTCAGTAACATCGGCCCTTCATCCGACTTA 3352
        |||||          |||||          |||||          |||||
Qy      21 ArgArgLeuGlnHisThr---PheValPheLeuArganPheSerIleuMetLeuLeu 38
        |||||          |||||          |||||          |||||
Db      3353 CATGACTGAGACACACTTGTCTTTTCTTTTCTATTCAGCCCGTGAATCTTTTA 3409
        |||||          |||||          |||||          |||||

RESULT 12
US-09-511-477-1
; Sequence 1, Application US/09511477
; Patent No. 6337202
; GENERAL INFORMATION:
; APPLICANT: JACOBSON, Myron K.
; APPLICANT: JACOBSON, Elaine L.
; APPLICANT: AME, Jean-Christophe
; APPLICANT: LIN, Winston
; TITLE OF INVENTION: GENES ENCODING SEVERAL POLY(ADP-RIBOSE) GLYCOPHYROLASE (PARG) ENZYME COMPOSITIONS AND ANTIBODIES IMMUNOREACTIVE
; TITLE OF INVENTION: THE PROTEINS AND FRAGMENTS THEREOF, AND ANTIBODIES IMMUNOREACTIVE
; FILE REFERENCE: NIND 201
; CURRENT APPLICATION NUMBER: US/09/511,477
; EARLIER FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 09/302,812
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 38
; SEQ ID NO 1
; LENGTH: 4070
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:

```

```

US-09-511-477-1
Alignment Scores:
Pred. No.: 80.7
Score: 62.50
Percent Similarity: 53.85%
Best Local Similarity: 43.59%
Query Match: 20.36%
DB: 3
Gaps: 1

US-10-031-158-14 (1-58) x US-09-511-477-1 (1-4070)
QY 1 MetGlnMetPheProPseProLeuPhePhePheLeuGlnLeuLeuYsgInSerSer 20
    |||||
Db 3293 ATGAGAGCCCTTCGCGCCCATGACGACTGTGACATCATCTGCGCTCCATCCGACTTA 3352
    |||||

QY 21 ArgArgLeuGlnIuHsIthr---PheValPheLeuArgAsnPheserLeuMetLeu 38
    |||||
Db 3353 CATGAGCTGAGACATACTTGTTCCTTTTTCCTTTCATTTCACGCCGATTCCTTTTA 3409
    |||||

RESULT 13
US-09-511-507-1
Sequence 1, Application US/09511507
Patent No. 6395543
GENERAL INFORMATION:
APPLICANT: JACOBSON, Myron K.
APPLICANT: JACOBSON, Elaine L.
APPLICANT: AME, Jean-Christophe
APPLICANT: LUV, Winston
TITLE OF INVENTION: GENES ENCODING SEVERAL POLY(ADP-RIBOSE) GLYCOHYDROLASE (PARG) EN
TITLE OF INVENTION: THE PROTEINS AND FRAGMENTS THEREOF, AND ANTIBODIES IMMUNOREACTIV
TITLE OF INVENTION: THEREMITH
FILE REFERENCE: N1AD 201
CURRENT APPLICATION NUMBER: US/09/511,507
CURRENT FILING DATE: 2000-02-23
PRIOR APPLICATION NUMBER: 09/302,812
PRIOR FILING DATE: 1999-04-30
NUMBER OF SEQ ID NOS: 38
SEQ ID NO 1
LENGTH: 4070
TYPE: DNA
ORGANISM: Bos taurus
FEATURE:
US-09-511-507-1

Alignment Scores:
Pred. No.: 80.7
Score: 62.50
Percent Similarity: 53.85%
Best Local Similarity: 43.59%
Query Match: 20.36%
DB: 3
Gaps: 1

US-10-031-158-14 (1-58) x US-09-511-507-1 (1-4070)
QY 1 MetGlnMetPheProPseProLeuPhePhePheLeuGlnLeuLeuYsgInSerSer 20
    |||||
Db 3293 ATGAGAGCCCTTCGCGCCCATGACGACTGTGACATCATCTGCGCTCCATCCGACTTA 3352
    |||||

QY 21 ArgArgLeuGlnIuHsIthr---PheValPheLeuArgAsnPheserLeuMetLeu 38
    |||||
Db 3353 CATGAGCTGAGACATACTTGTTCCTTTTTCCTTTCATTTCACGCCGATTCCTTTTA 3409
    |||||

RESULT 14
US-09-025-769B-274/c
Sequence 274, Application US/09025769B
Patent No. 6300064
GENERAL INFORMATION:
APPLICANT: Knappik, Achim
APPLICANT: Pack, Peter
APPLICANT: Illag, Vic
APPLICANT: Ge, Liming
APPLICANT: Moroney, Simon

```

APPLICANT: Plueckthun, Andreas  
TITLE OF INVENTION: Protein/(Poly)peptide libraries  
NUMBER OF SEQUENCES: 373  
CORRESPONDENCE ADDRESS:  
ADDRESSER: James F. Haley, Jr., Esq. c/o Fish & Neave  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10021  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/025,769B  
FILING DATE: 18-FEB-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 95 11 3021.0  
FILING DATE: 18-AUG-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: James F. Haley, Jr., Esq.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: MORPHO/5  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)596-9000  
TELEFAX: (212)596-9090  
INFORMATION FOR SEQ ID NO: 274:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2755 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: other nucleic acid  
DESCRIPTION: /desc = "synthetic vector"  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 3..509  
OTHER INFORMATION: /product= "gIIIp ss, myc tag, amber"  
FEATURE:  
NAME/KEY: CDS  
LOCATION: complement (1853..2509)  
OTHER INFORMATION: /product= "cat resistance"  
US-09-025-769B-274  
Alignment Scores:  
Pred. No.: 65.3 Length: 2755  
Score: 61.50 Matches: 15  
Percent Similarity: 52.83% Conservative: 13  
Best Local Similarity: 28.30% Mismatches: 8  
Query Match: 20.03% Indels: 17  
Gaps: 3  
US-10-031-158-14 (1-58) x US-09-025-769B-274 (1-2755)  
QY 2 GIMetPheProPseSerProLeuPhePhePheLeuInleuLysGlnSerSerArg 21  
DB 602 CAGGCCGCGCCCCCCCCCTTAAATTAAACGCGACAGACAAA 561  
QY 22 ArgLeuGluHisThrPheValPheLeuArgAsnPheSer----- 34  
DB 560 AAATATGCGACAACTCTCGGCATTTTTCACCTTCACAGTCAAGCTTATCA 510  
QY 35 ---LeuMetLeuLeuArgTyrTleGlyLysLysArgArg 46  
DB 509 AGACTCCTATTACGACAGTATGTTAGCAAAAGTGA AAAA 471  
RESULT 15  
US-09-490-070A-274/C  
Sequence 274, Application US/09490070A  
Patent No. 6696248

GENERAL INFORMATION:  
APPLICANT: Knappik, Achim  
Pach, Peter  
Ilag, Vic  
Ge, Liming  
McInerney, Simon  
Plueckthun, Andreas  
TITLE OF INVENTION: Protein/(Poly)peptide libraries  
NUMBER OF SEQUENCES: 373  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Colin G. Sandercock, Esq. c/o Heller Ehrman  
STREET: 1666 K Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/490,070A  
FILING DATE: 24-Jan-2000  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 95 11 3021.0  
FILING DATE: 18-AUG-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Colin G. Sandercock, Esq.  
REGISTRATION NUMBER: 31,298  
REFERENCE/DOCKET NUMBER: 37629-0005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 912-2000  
TELEFAX: (202) 912-2020  
INFORMATION FOR SEQ ID NO: 274:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2755 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: other nucleic acid  
DESCRIPTION: /desc = "synthetic vector"  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 3..509  
OTHER INFORMATION: /product= "gIIIp ss, myc tag, amber"  
FEATURE:  
NAME/KEY: CDS  
LOCATION: complement (1853..2509)  
OTHER INFORMATION: /product= "cat resistance"  
SEQUENCE DESCRIPTION: SEQ ID NO: 274:  
US-09-490-070A-274  
Alignment Scores:  
Pred. No.: 65.3 Length: 2755  
Score: 61.50 Matches: 15  
Percent Similarity: 52.83% Conservative: 13  
Best Local Similarity: 28.30% Mismatches: 8  
Query Match: 20.03% Indels: 17  
Gaps: 3  
US-10-031-158-14 (1-58) x US-09-490-070A-274 (1-2755)  
QY 2 GIMetPheProPseSerProLeuPhePhePheLeuInleuLysGlnSerSerArg 21  
DB 602 CAGGCCGCGCCCCCCCCCTTAAATTAAACGCGACAGACAAA 561  
QY 22 ArgLeuGluHisThrPheValPheLeuArgAsnPheSer----- 34  
DB 560 AAATATGCGACAACTCTCGGCATTTTTCACCTTCACAGTCAAGCTTATCA 510

```

Oy      5 ---LeumetleueuaqgyrrilegilylytalyAArgarg 46
Db      509 AGACTCTTATACGCAGTATGTTAGCAAACGTAGAAA 471

RESULT 16
US-09-490-153-274/c
; Sequence 274, Application US/09490153
Patent No. 6706484
GENERAL INFORMATION:
APPLICANT: Knappik, Achim
Pack, Peter
ilag, Vic
Ge, Liming
Moroney, Simon
Pluckeltnun, Andreas
TITLE OF INVENTION: Protein/(Poly)peptide libraries
NUMBER OF SEQUENCES: 373
CORRESPONDENCE ADDRESS:
ADDRESSEE: James F. Haley, Jr., Esq. c/o Fish & Neave
STREET: 151 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10021
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (ERO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/490,153
FILING DATE: 24-Jan-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/025,769B
FILING DATE: 18-FEB-1998
APPLICATION NUMBER: EP 95 11 3021.0
FILING DATE: 18-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: James F. Haley, Jr., Esq.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: MORPHO/S
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)596-9000
TELEFAX: (212)596-9090
INFORMATION FOR SEQ ID NO: 274:
SEQUENCE CHARACTERISTICS:
LENGTH: 2755 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "synthetic vector"
FEATURE:
NAME/KEY: CDS
LOCATION: 3..509
OTHER INFORMATION: /product= "giilp ss, myc tag, amber"
codon"
US-09-490-153-274
Alignment Scores:
Pred. No.: 65.3 Length: 2755
Score: 61.50 Matches: 15
Percent Similarity: 52.83% Conservative: 13
Best Local Similarity: 28.30% Mismatches: 8
Query Match: 20.03% Indels: 17
DB: Gaps: 3
US-10-031-158-14 (1-58) x US-09-490-153-274 (1-2755)

```

```

Oy      2   GlnMetPheProSerProLeuNhePheLeuGlnLeuLysGlnSerArg 2
          ||| ||||| ||||| ||| ||||| ||||| ||||| ||||| |||||
Db      602 CAGCGGGCCCCCCCCCTT-----TTAATTAAACGCAGCAAAA 561
Oy      22   ArgLeuGlnIsthrPheValPheLeuArgAsnPheser----- 34
          :||::||| ||||| ||||| ||||| ||||| ||||| |||||
Db      560 AAAATGTGCGCACAT-----CTGGGCCATTTTTTCATCTCACAAGTCAGCTTATCA 510
Oy      35   --LeuMetLeuLeuArgTyrlleGlyLysValArgArg 46
          ||||::||| ::||::||| ||||| ||||| ||||| |||||
Db      509 AGACTCTTATTACGACAGTAGTATTGACAAACGTAGAAAA 471

RESULT 17
US-09-250-609-79
; Sequence 79, Application US/09250609A
; Patent No. 6458943
; GENERAL INFORMATION:
; APPLICANT: Byrne, Jennifer A.
; TITLE OF INVENTION: Members of the D52 Gene Family
; FILE REFERENCE: 1383.0210002
; CURRENT APPLICATION NUMBER: US/09/250,609A
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 79
; LENGTH: 339
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-250-609-79

Alignment Scores:
Pred. No.:           3.99           Length:           339
Score:              61.00           Matches:         18
Percent Similarity: 48.94%           Conservative:     5
Best Local Similarity: 38.30%        Mismatches:       14
Query Match:        19.87%           Indels:          10
                               Gaps:            2
                               4

US-10-031-158-14 (1-58) x US-09-250-609-79 (1-339)

Oy      13   LeuGlnLeuLeuLys-----GlnSerSerArgLeuGlnIsthrPheValPheLeu 30
          ||| ||||| ||||| ||| ||||| ||||| ||||| |||||
Db      197 CTCGAGTTGTAGAAGCGCTGCACAGAGGGGAGAGAGATT----- 235
Oy      31   ArgAsnPheSerLeuMetLeuLeuArgTyrlleGlyLysValArgAlaThrArgPhe 50
          ::||::||| ||||| ||||| ||||| ||||| ||||| |||||
Db      236 ---AGCTTCGGGCTGAGCTTGCTTAAGGTGAAGAAATTTGCACTGCGCCAGGTGC 292
Oy      51   TrpAspProArgArgGLYThr 57
          ||| ||||| ||||| ||||| ||||| ||||| |||||
Db      293 TGGCAGCCCAAAGAGAGGCACT 313

RESULT 18
US-09-250-611-79
; Sequence 79, Application US/09250611
; Patent No. 6528283
; GENERAL INFORMATION:
; APPLICANT: Byrne, Jennifer A.
; TITLE OF INVENTION: Members of the D52 Gene Family
; FILE REFERENCE: 1383.0210001
; CURRENT APPLICATION NUMBER: US/09/250,611
; CURRENT FILING DATE: 1999-02-17
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 79
; LENGTH: 339
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-250-611-79

Alignment Scores:
```

Pred. No.: 3.99 Length: 339  
Score: 61.00 Matches: 18  
Percent Similarity: 48.94% Conservative: 5  
Best Local Similarity: 38.30% Mismatches: 14  
Query Match: 19.87% Indels: 10  
DB: 4 Gaps: 2

US-10-031-158-14 (1-58) x US-09-250-611-79 (1-339)

QY 13 LeuGlnLeuLeuLys-----GlnSerSerArgArgLeuGlnHisThrPheValPheLeu 30

DB 197 CTCGACGTTAGAGGCTGACAGGGGAGGAGAAATT----- 235

QY 31 ArgAsnPheserLeuMetLeuLeuArgTyrIleGlyLysLysArgArgAlaThrArgPhe 50

DB 236 ---AGCTTGGGCTGAGCTTGCTAAGCTGGAAGAAGAAATTGCACTGCGCCGAGGTGC 292

QY 51 TTPAspProArgArgGlyThr 57

DB 293 TGGCAGCCCAAGAGAGGCACT 313

RESULT 19

US-09-250-609-91  
Sequence 91, Application US/09250609A  
Patent No. 6458943

GENERAL INFORMATION:

APPLICANT: Byrne, Jennifer A.  
TITLE OF INVENTION: Members of the D52 Gene Family

FILE REFERENCE: 1383.0210002  
CURRENT APPLICATION NUMBER: US/09/250,609A

CURRENT FILING DATE: 1999-02-17  
NUMBER OF SEQ ID NOS: 108

SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 91

LENGTH: 399  
TYPE: DNA

ORGANISM: Homo sapiens  
US-09-250-609-91

Alignment Scores:

Pred. No.: 5.02 Length: 399  
Score: 61.00 Matches: 12  
Percent Similarity: 59.26% Conservative: 4  
Best Local Similarity: 44.44% Mismatches: 11  
Query Match: 19.87% Indels: 0  
DB: 4 Gaps: 0

US-10-031-158-14 (1-58) x US-09-250-609-91 (1-399)

QY 31 ArgAsnPheserLeuMetLeuLeuArgTyrIleGlyLysLysArgArgAlaThrArgPhe 50

DB 227 AAGAGCTTGCGGCTGAGCTTGCTAAGCTGGAAGAAGAAATTGCACTGCGCCGAGGTGC 286

QY 51 TTPAspProArgArgGlyThr 57

DB 287 TGGCAGCCCAAGAGAGGCACT 307

RESULT 20

US-09-250-611-91  
Sequence 91, Application US/09250611  
Patent No. 6528283

GENERAL INFORMATION:

APPLICANT: Byrne, Jennifer A.  
TITLE OF INVENTION: Members of the D52 Gene Family

FILE REFERENCE: 1383.0210001  
CURRENT APPLICATION NUMBER: US/09/250,611

CURRENT FILING DATE: 1999-02-17  
NUMBER OF SEQ ID NOS: 108  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 91

LENGTH: 399  
TYPE: DNA

ORGANISM: Homo sapiens  
US-09-250-611-91

Alignment Scores:

Pred. No.: 5.02 Length: 399  
Score: 61.00 Matches: 12  
Percent Similarity: 59.26% Conservative: 4  
Best Local Similarity: 44.44% Mismatches: 11  
Query Match: 19.87% Indels: 0  
DB: 4 Gaps: 0

US-10-031-158-14 (1-58) x US-09-250-611-91 (1-399)

QY 31 ArgAsnPheserLeuMetLeuLeuArgTyrIleGlyLysLysArgArgAlaThrArgPhe 50

DB 227 AAGAGCTTGCGGCTGAGCTTGCTAAGCTGGAAGAAGAAATTGCACTGCGCCGAGGTGC 286

QY 51 TTPAspProArgArgGlyThr 57

DB 287 TGGCAGCCCAAGAGAGGCACT 307

RESULT 21

US-09-250-609-89  
Sequence 89, Application US/09250609A  
Patent No. 6458943

GENERAL INFORMATION:  
APPLICANT: Byrne, Jennifer A.  
TITLE OF INVENTION: Members of the D52 Gene Family

FILE REFERENCE: 1383.0210002  
CURRENT APPLICATION NUMBER: US/09/250,609A

CURRENT FILING DATE: 1999-02-17  
NUMBER OF SEQ ID NOS: 108

SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 89

LENGTH: 420  
TYPE: DNA

ORGANISM: Homo sapiens  
US-09-250-609-89

Alignment Scores:

Pred. No.: 5.4 Length: 420  
Score: 61.00 Matches: 12  
Percent Similarity: 59.26% Conservative: 4  
Best Local Similarity: 44.44% Mismatches: 11  
Query Match: 19.87% Indels: 0  
DB: 4 Gaps: 0

US-10-031-158-14 (1-58) x US-09-250-609-89 (1-420)

QY 31 ArgAsnPheserLeuMetLeuLeuArgTyrIleGlyLysLysArgArgAlaThrArgPhe 50

DB 230 AAGAGCTTGCGGCTGAGCTTGCTAAGCTGGAAGAAGAAATTGCACTGCGCCGAGGTGC 289

QY 51 TTPAspProArgArgGlyThr 57

DB 290 TGGCAGCCCAAGAGAGGCACT 310

RESULT 22

US-09-250-611-89  
Sequence 89, Application US/09250611  
Patent No. 6528283

GENERAL INFORMATION:

APPLICANT: Byrne, Jennifer A.  
TITLE OF INVENTION: Members of the D52 Gene Family

FILE REFERENCE: 1383.0210001  
CURRENT APPLICATION NUMBER: US/09/250,611

CURRENT FILING DATE: 1999-02-17  
NUMBER OF SEQ ID NOS: 108  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 89

LENGTH: 420  
TYPE: DNA

```
/ ORGANISM: Homo sapiens
US-09-250-611-89

Alignment Scores:
Pred. No.: 5.4 Length: 420
Score: 61.00 Matches: 12
Percent Similarity: 59.26% Conservative: 4
Best Local Similarity: 44.44% Mismatches: 11
Query Match: 19.87% Indels: 0
DB: 4 Gaps: 0

US-10-031-158-14 (1-58) x US-09-250-611-89 (1-420)

QY 31 ArgAnpHeSerLeuMeLeuLeuArgTyrILlegLYLyLybArgArgAlaThrArgPhe 50
DB 230 AAGAGCTTCGGGCTGAGCTTGCTTAAGGTGAAGAAATTGTCACTCTGCCCGCAGGTGC 289

QY 51 TTPAaPProARgArgGLyThr 57
DB 290 TGGCAGCCAAAGAGAGGCACT 310

RESULT 23
US-09-250-609-77
; Sequence 77, Application US/09250609A
; Patent No. 6458943
; GENERAL INFORMATION:
; APPLICANT: Byrne, Jennifer A.
; TITLE OF INVENTION: Members of the D52 Gene Family
; FILE REFERENCE: 1383.0210002
; CURRENT APPLICATION NUMBER: US/09/250,609A
; CURRENT FILING DATE: 1999-02-17
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 77
; LENGTH: 424
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-250-609-77

Alignment Scores:
Pred. No.: 5.47 Length: 424
Score: 61.00 Matches: 12
Percent Similarity: 59.26% Conservative: 4
Best Local Similarity: 44.44% Mismatches: 11
Query Match: 19.87% Indels: 0
DB: 4 Gaps: 0

US-10-031-158-14 (1-58) x US-09-250-609-77 (1-424)

QY 31 ArgAnpHeSerLeuMeLeuLeuArgTyrILlegLYLyLybArgArgAlaThrArgPhe 50
DB 239 AAGAGCTTCGGGCTGAGCTTGCTTAAGGTGAAGAAATTGTCACTCTGCCCGCAGGTGC 298

QY 51 TTPAaPProARgArgGLyThr 57
DB 299 TGGCAGCCAAAGAGAGGCACT 319

RESULT 24
US-09-250-611-77
; Sequence 77, Application US/09250611
; Patent No. 6528283
; GENERAL INFORMATION:
; APPLICANT: Byrne, Jennifer A.
; TITLE OF INVENTION: Members of the D52 Gene Family
; FILE REFERENCE: 1383.0210001
; CURRENT APPLICATION NUMBER: US/09/250,611
; CURRENT FILING DATE: 1999-02-17
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 77
; LENGTH: 424
; TYPE: DNA
```

```
/ ORGANISM: Homo sapiens
US-09-250-611-77

Alignment Scores:
Pred. No.: 5.47 Length: 424
Score: 61.00 Matches: 12
Percent Similarity: 59.26% Conservative: 4
Best Local Similarity: 44.44% Mismatches: 11
Query Match: 19.87% Indels: 0
DB: 4 Gaps: 0

US-10-031-158-14 (1-58) x US-09-250-611-77 (1-424)

QY 31 ArgAnpHeSerLeuMeLeuLeuArgTyrILlegLYLyLybArgArgAlaThrArgPhe 50
DB 239 AAGAGCTTCGGGCTGAGCTTGCTTAAGGTGAAGAAATTGTCACTCTGCCCGCAGGTGC 298

QY 51 TTPAaPProARgArgGLyThr 57
DB 299 TGGCAGCCAAAGAGAGGCACT 319

RESULT 25
US-09-250-609-90
; Sequence 90, Application US/09250609A
; Patent No. 6458943
; GENERAL INFORMATION:
; APPLICANT: Byrne, Jennifer A.
; TITLE OF INVENTION: Members of the D52 Gene Family
; FILE REFERENCE: 1383.0210002
; CURRENT APPLICATION NUMBER: US/09/250,609A
; CURRENT FILING DATE: 1999-02-17
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 90
; LENGTH: 471
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-250-609-90

Alignment Scores:
Pred. No.: 6.35 Length: 471
Score: 61.00 Matches: 12
Percent Similarity: 59.26% Conservative: 4
Best Local Similarity: 44.44% Mismatches: 11
Query Match: 19.87% Indels: 0
DB: 4 Gaps: 0

US-10-031-158-14 (1-58) x US-09-250-609-90 (1-471)

QY 31 ArgAnpHeSerLeuMeLeuLeuArgTyrILlegLYLyLybArgArgAlaThrArgPhe 50
DB 248 AAGAGCTTCGGGCTGAGCTTGCTTAAGGTGAAGAAATTGTCACTCTGCCCGCAGGTGC 307

QY 51 TTPAaPProARgArgGLyThr 57
DB 308 TGGCAGCCAAAGAGAGGCACT 328

RESULT 26
US-09-250-611-90
; Sequence 90, Application US/09250611
; Patent No. 6528283
; GENERAL INFORMATION:
; APPLICANT: Byrne, Jennifer A.
; TITLE OF INVENTION: Members of the D52 Gene Family
; FILE REFERENCE: 1383.0210001
; CURRENT APPLICATION NUMBER: US/09/250,611
; CURRENT FILING DATE: 1999-02-17
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 90
; LENGTH: 471
; TYPE: DNA
```

```

; ORGANISM: Homo sapiens
US-09-250-611-90

Alignment Scores:
Pred. No.:      6.35          Length:    471
Score:         61.00        Matches:     12
Percent Similarity: 59.26%   Conservative: 4
Best Local Similarity: 44.44% Mismatches:  11
Query Match:    19.87%       Indels:      0
DB:             4           Gaps:         0

US-10-031-158-14 (1-58) x US-09-250-611-90 (1-471)

QY      31 ArgAenPhSeSerLeuMetLeuLauArgTYrILeGlyLyIySArGaRgaLatMrARgPhe 50
       ::::||||| |::: ||||| ::|||::: |||||
Db      248 AAGACCTTCGGCGCTGCATGTCTTAAGTGAAGAAGAAATTGTCACTTGCCGCAAGTGC 307
               |||:::||:::|||::|||

QY      51 TTPAAPProARGArgGLyThr 57
       ||| |||:::|||::|||
Db      308 TGCGACGCCAAGAGAGGCACT 328
               |||:::|||::|||

RESULT 27
US-09-250-609-95
; Sequence 95, Application US/09250609A
; Patent No. 6458943
; GENERAL INFORMATION:
; APPLICANT: Byrne, Jennifer A.
; TITLE OF INVENTION: Members of the D52 Gene Family
; FILE REFERENCE: 1383.0210002
; CURRENT APPLICATION NUMBER: US/09/250,609A
; CURRENT FILING DATE: 1999-02-17
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 95
; LENGTH: 540
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-250-609-95

Alignment Scores:
Pred. No.:      7.71          Length:    540
Score:         61.00        Matches:     12
Percent Similarity: 59.26%   Conservative: 4
Best Local Similarity: 44.44% Mismatches:  11
Query Match:    19.87%       Indels:      0
DB:             4           Gaps:         0

US-10-031-158-14 (1-58) x US-09-250-609-95 (1-540)

QY      31 ArgAenPhSeSerLeuMetLeuLauArgTYrILeGlyLyIySArGaRgaLatMrARgPhe 50
       ::::||||| |::: ||||| ::|||::: |||||
Db      96 AAGACTTCGGCGCTGCATGTCTTAAGTGAAGAAGAAATTGTCACTTGCCGCAAGTGC 155
               |||:::|||:::|||::|||

QY      51 TTPAAPProARGArgGLyThr 57
       ||| |||:::|||::|||
Db      156 TGCGACGCCAAGAGAGGCACT 176
               |||:::|||::|||

RESULT 28
US-09-250-611-95
; Sequence 95, Application US/09250611
; Patent No. 6528283
; GENERAL INFORMATION:
; APPLICANT: Byrne, Jennifer A.
; APPLICANT: Bassec, Paul
; TITLE OF INVENTION: Members of the D52 Gene Family
; FILE REFERENCE: 1383.0210001
; CURRENT APPLICATION NUMBER: US/09/250,611
; CURRENT FILING DATE: 1999-02-17
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 95
; LENGTH: 540
; TYPE: DNA
```

```

; ORGANISM: Homo sapiens
US-09-250-611-95

Alignment Scores:
Pred. No.:          7.71      Length:          540
Score:              61.00     Matches:         12
Percent Similarity: 59.26%    Conservative:   4
Best Local Similarity: 44.44% Mismatches:        11
Query Match:       19.87%    Indels:           0
DB:                4        Gaps:            0

US-10-031-158-14 (1-58) x US-09-250-611-95 (1-540)
QY      31 ArgAenPheserLeuMeleuRphneUgTYrileGlyLyelVysArgArGaLaThraRgphe 50
      ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||
Db      96 AAGACCTTCGGGCTTAGCTTGCTTAAGGTGAAGAAGAAATTGTCACTTCCGCCAGGTGC 155

QY      51 TTPaePRoARgArGLyThr 57
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||
Db      156 TCGCAGCCCAAGAGAGGCACT 176

RESULT 29
US-09-495-050A-103/c
; Sequence 103, Application US/09495050A
; Patent No. 6492505
; GENERAL INFORMATION:
; APPLICANT: Roopa, Reddy
; APPLICANT: Gugler, Karl, J.
; APPLICANT: Au-Young, Janice
; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-AS
; FILE REFERENCE: PA-0013 US
; CURRENT APPLICATION NUMBER: US/09/495,050A
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/118,318
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: PERL Program
; SEQ ID NO 103
; LENGTH: 836
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6492505 1331917CTL
US-09-495-050A-103

Alignment Scores:
Pred. No.:          14.3      Length:          836
Score:              61.00     Matches:         17
Percent Similarity: 65.52%    Conservative:   2
Best Local Similarity: 58.62% Mismatches:         9
Query Match:       19.87%    Indels:           1
DB:                4        Gaps:            0

US-10-031-158-14 (1-58) x US-09-495-050A-103 (1-836)
QY      4 PhePrioProSePrOleuPhephePhe-LeuGlnLeuLeuLysGlnSerseRgArGLE 23
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||
Db      474 TTCTGCCGCTCTCTCTCTCTCTTTGCGCCTTCTCCAACTCTCTGAGACGCT 415

QY      23 uGUHstHrPhValPheLeuArg 31
      ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||
Db      414 CTTTCTCTCTGTTCTCTCTCTGCG 390

RESULT 30
US-09-587-184-1/c
; Sequence 1, Application US/09587184
; Patent No. 6479729
; GENERAL INFORMATION:
; APPLICANT: Camprochiaro, Peter A.
; APPLICANT: Zack, Donald J.
; TITLE OF INVENTION: Mouse Model for Ocular Neovascularization
; FILE REFERENCE: 001107.00011

```

```
/ CURRENT APPLICATION NUMBER: US/09/587,184
/ CURRENT FILING DATE: 2000-06-05
/ NUMBER OF SEQ ID NOS: 7
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 1
/ LENGTH: 2136
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: VEGF construct
US-09-587-184-1

Alignment Scores:
Pred. No.: 54.1 Length: 2136
Score: 61.00 Matches: 15
Percent Similarity: 61.11% Conservative: 7
Best Local Similarity: 41.67% Mismatches: 12
Query Match: 19.87% Indels: 2
DB: 4 Gaps: 1

US-10-031-158-14 (1-58) x US-09-587-184-1 (1-2136)

QY 6 ProSerProLeuPhePheLeuGlnLeuLysGlnSerSerArg-----ArgLeu 23
Db 1730 CCCAGCCCGCCCTCTACTTCTGCGCTCTCGCGGCGAGCATGCGCTCTCGGCTG 1671
QY 24 GlnHisThrPheValPheLeuArgPheSerLeuMetLeuLeuArg 39
Db 1670 CGACATCTTCTGCGTGGAGCGGCGATCTGCTCTGCTTCTGCGG 1623

RESULT 31
US-09-688-188B-9/c
/ Sequence 9, Application US/09688188B
/ Patent No. 6656716
/ GENERAL INFORMATION:
/ APPLICANT: PLOWMAN, GREGORY
/ APPLICANT: MARTINEZ, RICARDO
/ APPLICANT: WHYTE, DAVID
/ TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
/ FILE REFERENCE: 038602/0328
/ CURRENT APPLICATION NUMBER: US/09/688,188B
/ CURRENT FILING DATE: 2000-10-16
/ PRIOR APPLICATION NUMBER: 09/291,417
/ PRIOR FILING DATE: 1999-04-14
/ PRIOR APPLICATION NUMBER: 60/081,784
/ NUMBER OF SEQ ID NOS: 155
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 9
/ LENGTH: 3798
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-688-188B-9

Alignment Scores:
Pred. No.: 122 Length: 3798
Score: 61.00 Matches: 17
Percent Similarity: 65.52% Conservative: 2
Best Local Similarity: 58.62% Mismatches: 9
Query Match: 19.87% Indels: 1
DB: 4 Gaps: 0

US-10-031-158-14 (1-58) x US-09-688-188B-9 (1-3798)

QY 4 PheProProSerProLeuPhePhePhe-LeuGlnLeuLysGlnSerSerArgArg 23
Db 1371 TTCTGCGCGCTCTCTCTCTCTCTTCTTGGCGCTTCTCTCACTCTCTAGAGCCT 1312
QY 23 uGlnHisThrPheValPheLeuArg 31
Db 1311 CTTTCTCTTGTCTCTCTCTGCG 1287

RESULT 32
```

```
US-09-291-417D-9/c
/ Sequence 9, Application US/09291417D
/ Patent No. 6680170
/ GENERAL INFORMATION:
/ APPLICANT: PLOWMAN, GREGORY
/ APPLICANT: MARTINEZ, RICARDO
/ APPLICANT: WHYTE, DAVID
/ TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
/ FILE REFERENCE: 038602/0329
/ CURRENT APPLICATION NUMBER: US/09/291,417D
/ CURRENT FILING DATE: 1999-04-13
/ PRIOR APPLICATION NUMBER: 60/081,784
/ PRIOR FILING DATE: 1998-04-14
/ NUMBER OF SEQ ID NOS: 155
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 9
/ LENGTH: 3798
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-291-417D-9

Alignment Scores:
Pred. No.: 122 Length: 3798
Score: 61.00 Matches: 17
Percent Similarity: 65.52% Conservative: 2
Best Local Similarity: 58.62% Mismatches: 9
Query Match: 19.87% Indels: 1
DB: 4 Gaps: 0

US-10-031-158-14 (1-58) x US-09-291-417D-9 (1-3798)

QY 4 PheProProSerProLeuPhePhePhe-LeuGlnLeuLysGlnSerSerArgArg 23
Db 1371 TTGCGCGCTCTCTCTCTCTCTTCTTGGCGCTTCTCTCAACTCTCTAGAGCCT 1312
QY 23 uGlnHisThrPheValPheLeuArg 31
Db 1311 CTTTCTCTTGTCTCTCTCTGCG 1287

RESULT 33
US-09-651-011A-3/c
/ Sequence 3, Application US/09651011A
/ Patent No. 6346416
/ GENERAL INFORMATION:
/ APPLICANT: Nicholas M. Dean
/ APPLICANT: Lex M. Cowsett
/ TITLE OF INVENTION: ANTISENSE MODULATION OF HPK/GCK-LIKE KINASE EXPRESSION
/ FILE REFERENCE: RTS-0168
/ CURRENT APPLICATION NUMBER: US/09/651,011A
/ CURRENT FILING DATE: 2000-08-29
/ NUMBER OF SEQ ID NOS: 49
/ SEQ ID NO 3
/ LENGTH: 4266
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ NAME/KEY: CDS
/ LOCATION: (1)...(3528)
US-09-651-011A-3

Alignment Scores:
Pred. No.: 144 Length: 4266
Score: 61.00 Matches: 17
Percent Similarity: 65.52% Conservative: 2
Best Local Similarity: 58.62% Mismatches: 9
Query Match: 19.87% Indels: 1
DB: 3 Gaps: 0

US-10-031-158-14 (1-58) x US-09-651-011A-3 (1-4266)

QY 4 PheProProSerProLeuPhePhePhe-LeuGlnLeuLysGlnSerSerArgArg 23
Db 1254 TTGCGCGCTCTCTCTCTCTCTTCTTGGCGCTTCTCTCAACTCTCTAGAGCCT 1195
```



RESULT 37  
US-08-943-731-7/c  
; Sequence 7, Application US/08943731  
; Patent No. 6265157  
; GENERAL INFORMATION:  
; APPLICANT: PROCKOP, DARWIN J.  
; APPLICANT: SPOTILIA, LORETTA D.  
; APPLICANT: DELTAS, CONSTANTINOS D.  
; APPLICANT: SEREDA, LARISA  
; APPLICANT: LARSON, ANDREA W.  
; APPLICANT: PACK, MICHAEL  
; APPLICANT: COLIGE, ALAIN  
; APPLICANT: EARLY, JAMES  
; APPLICANT: KORKKO, JARMO  
; APPLICANT: ALA-KORKO, LEENA, et al.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING  
; TITLE OF INVENTION: ALTERED TYPE I OR TYPE IX COLLAGEN GENE SEQUENCES  
; NUMBER OF SEQUENCES: 666  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.  
; STREET: ONE COMMERCE SQUARE, 2005 MARKET STREET, 22ND  
; STREET: FLR.  
; CITY: PHILADELPHIA  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103-7086  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/943,731  
; FILING DATE: 03-OCT-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/212,322  
; FILING DATE: 14-MAR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/803,628  
; FILING DATE: 03-DEC-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DOYLE LEARY Ph.D., KATHRYN  
; REGISTRATION NUMBER: 36,317  
; REFERENCE/DOCKET NUMBER: 9598-27  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-965-1284  
; TELEFAX: 215-567-2991  
; TELEX: 831-494  
; INFORMATION FOR SEQ. ID NO. 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1463 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-943-731-7  
  
Alignment Scores:  
Pred. No.: 37.5 Length: 1463  
Score: 60.50 Matches: 17  
Percent Similarity: 42.86% Conservative: 4  
Best Local Similarity: 34.69% Mismatches: 15  
Query Match: 19.71% Indels: 13  
DB: 3 Gaps: 1  
  
US-10-031-158-14 (1-58) X US-08-943-731-7 (1-1463)  
QY 4 PhePProPSePProLeuPhePheLeuGlnLeuLeuLysGlnSerSerArgArgLeu 23  
DB 1366 TTCGCCGCCCTCCCTCTTCTTCTTCAATCTTAAAGCTCGCTCCTCAT 1307  
QY 24 GIWHLS-----ThrPheValPheLeu 30

DB 1306 CAGCGCGGTGACAGCGCGGAGCGCTGCGCAGGAGCTGTACTCTTTCAGTCTC 1247  
QY 3i ArgPhePheSerLeuMetLeuLeuArg 39  
DB 1246 AGGAATTTAACAAAGCTTAGTCCG 1220  
  
RESULT 38  
US-08-943-731-1/c  
; Sequence 1, Application US/08943731  
; Patent No. 6265157  
; GENERAL INFORMATION:  
; APPLICANT: PROCKOP, DARWIN J.  
; APPLICANT: SPOTILIA, LORETTA D.  
; APPLICANT: DELTAS, CONSTANTINOS D.  
; APPLICANT: SEREDA, LARISA  
; APPLICANT: LARSON, ANDREA W.  
; APPLICANT: PACK, MICHAEL  
; APPLICANT: COLIGE, ALAIN  
; APPLICANT: EARLY, JAMES  
; APPLICANT: KORKKO, JARMO  
; APPLICANT: ALA-KORKO, LEENA, et al.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING  
; TITLE OF INVENTION: ALTERED TYPE I OR TYPE IX COLLAGEN GENE SEQUENCES  
; NUMBER OF SEQUENCES: 666  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.  
; STREET: ONE COMMERCE SQUARE, 2005 MARKET STREET, 22ND  
; STREET: FLR.  
; CITY: PHILADELPHIA  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103-7086  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/943,731  
; FILING DATE: 03-OCT-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/212,322  
; FILING DATE: 14-MAR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/803,628  
; FILING DATE: 03-DEC-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DOYLE LEARY Ph.D., KATHRYN  
; REGISTRATION NUMBER: 36,317  
; REFERENCE/DOCKET NUMBER: 9598-27  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-965-1284  
; TELEFAX: 215-567-2991  
; TELEX: 831-494  
; INFORMATION FOR SEQ. ID NO. 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 18609 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-943-731-1  
  
Alignment Scores:  
Pred. No.: 1.38e+03 Length: 18609  
Score: 60.50 Matches: 17  
Percent Similarity: 42.86% Conservative: 4  
Best Local Similarity: 34.69% Mismatches: 15  
Query Match: 19.71% Indels: 13  
DB: 3 Gaps: 1



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 8, 2004, 08:54:21 ; Search time 271 Seconds  
(without alignments)  
1180.016 Million cell updates/sec

Title: US-10-031-158-14

Perfect score: 307  
Sequence: 1 MOMPSPPLFFELQLLKSS.....RYGKKRRATRPWDRGRGP 58

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4085731 seqs, 2756760397 residues

Total number of hits satisfying chosen parameters: 8171462

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame\_p2n.model -DEV=x1h  
-O=/cg2\_1/USPTO\_EPOOL/US10031158/runat\_06122004\_082701\_16006/app\_query.fasta\_1.199  
-DB=Published Applications NA -QMT=fastcap -SUFFIX=trpb -MINMATCH=0.1  
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=blousum62  
-TRANS=human40.cdi -LIST=45 -LOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=40 -MODE=local -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USPR=US10031158\_0CGN\_1.1354\_0runat\_06122004\_082701\_16006  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:\*  
1: /cg2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cg2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cg2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cg2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cg2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cg2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cg2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cg2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cg2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
10: /cg2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cg2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cg2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cg2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cg2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cg2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
16: /cg2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*  
17: /cg2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*  
18: /cg2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
19: /cg2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*  
20: /cg2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
21: /cg2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Prod. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

\*

Result No.	Score	Query Match	Length	DB ID	Description
1	307	100.0	470	9 US-09-864-761-15428	Sequence 15428, A
2	307	100.0	477	9 US-09-864-761-15438	Sequence 15438, A
3	307	100.0	825	15 US-10-340-536-29	Sequence 29, Appl
4	307	100.0	1027	15 US-10-205-823-40	Sequence 404, App
5	307	100.0	1395	18 US-10-643-755A-23	Sequence 23, Appl
6	307	100.0	1418	10 US-09-957-708-16	Sequence 16, Appl
7	302	98.4	316	9 US-09-864-761-31943	Sequence 31943, A
8	294	95.8	1421	9 US-09-954-456-317	Sequence 317, App
9	294	95.8	1799	15 US-10-101-510-492	Sequence 492, App
10	284	92.5	1586	10 US-09-960-706-576	Sequence 676, App
11	284	92.5	1586	10 US-09-873-319-424	Sequence 424, App
12	245	79.8	1155	9 US-09-925-300-58	Sequence 58, Appl
13	75	24.4	1813	17 US-10-437-963-35632	Sequence 35632, A
14	73.5	23.9	618	13 US-10-027-632-136731	Sequence 136731, A
15	73.5	23.9	618	13 US-10-027-632-136731	Sequence 136731, A
16	72.5	23.6	618	13 US-10-027-632-105504	Sequence 105504, A
17	72.5	23.6	618	13 US-10-027-632-105505	Sequence 105505, A
18	72.5	23.6	618	13 US-10-027-632-105505	Sequence 105504, A
19	72.5	23.6	618	15 US-10-027-632-105505	Sequence 105505, A
20	72.5	23.6	27118	10 US-09-764-891-10230	Sequence 10230, A
21	72.5	23.6	27118	15 US-10-205-428-1018	Sequence 63943, A
22	71.5	23.3	328	16 US-10-424-599-63943	Sequence 136730, A
23	70.5	23.0	618	13 US-10-027-632-136730	Sequence 136730, A
24	70.5	22.0	618	15 US-10-027-632-136730	Sequence 127254, A
25	69.5	22.6	353	16 US-10-424-599-127254	Sequence 127254, A
26	69.5	22.6	2145	16 US-10-282-122A-22460	Sequence 22460, A
27	69.5	22.6	2241	9 US-09-881-752A-321	Sequence 321, App
28	68.5	22.3	585	13 US-10-027-632-203503	Sequence 203503, A
29	68.5	22.3	585	15 US-10-027-632-203503	Sequence 203503, A
30	67.5	22.0	535	18 US-10-363-345A-15079	Sequence 15079, A
31	67.5	22.0	535	18 US-10-363-345A-15080	Sequence 15080, A
32	67.5	22.0	942	9 US-09-962-833-256	Sequence 256, App
33	67.5	22.0	1037	15 US-10-172-118-1338	Sequence 1338, App
34	67.5	22.0	1037	16 US-10-342-887-1338	Sequence 1338, App
35	67.5	22.0	2532	10 US-09-918-648B-4	Sequence 4, Appl
36	67.5	22.0	11838	10 US-09-764-891-7020	Sequence 7020, App
37	67.5	22.0	13814	15 US-10-311-455-1165	Sequence 1165, App
38	67	21.8	314	18 US-10-425-115-4936	Sequence 25580, A
39	67	21.8	1265	13 US-10-027-632-252580	Sequence 252580, A
40	67	21.8	1265	15 US-10-027-632-252580	Sequence 252580, A
41	67	21.8	1726	18 US-10-425-115-128054	Sequence 128054, A
42	67	21.8	29829	13 US-10-087-192-694	Sequence 694, App
43	67	21.7	328	15 US-10-312-841-2	Sequence 2, Appl
44	66.5	21.7	1316	18 US-10-425-115-135105	Sequence 135105, A
45	66.5	21.7	1316	18 US-10-425-115-135105	Sequence 135105, A

#### ALIGNMENTS

RESULT 1  
US-09-864-761-15428  
Sequence 15428, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aecm1ca-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
PRIORITY FILING DATE: 2001-05-23  
PRIORITY FILING DATE: 2000-02-04  
PRIORITY FILING DATE: 2000-05-26  
PRIORITY FILING DATE: 2000-05-26  
PRIORITY FILING DATE: 2000-08-03  
PRIORITY APPLICATION NUMBER: US 60/180,312  
PRIORITY APPLICATION NUMBER: US 60/207,456  
PRIORITY APPLICATION NUMBER: US 09/632,366  
PRIORITY APPLICATION NUMBER: GB 24263,6

```
/ PRIOR FILING DATE: 2000-10-04
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 2000-09-27
/ PRIOR APPLICATION NUMBER: PCT/US01/00666
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00667
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00664
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00669
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00665
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00668
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00663
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00662
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00661
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00670
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: US 09/774,203
/ PRIOR FILING DATE: 2001-01-29
/ NUMBER OF SEQ ID NOS: 49117
/ SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 15428
/ LENGTH: 470
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AF159056.1
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.2
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
US-09-864-761-15428

Alignment Scores:
Pred. No.: 5,1e-36 Length: 470
Score: 307.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-031-158-14 (1-58) x US-09-864-761-15428 (1-470)

QY 1 MetGlnMetPheProPsePProLeuPhePheLeuGlnLeuLeuYsgInSerSer 20
DB 203 ATGCAAGATGTTTCCCAAGCCCACTATTTTCTTCAATTGCTGAAACAAAGCTCC 262
QY 21 ArgArgLeuGlnIuIstIrhPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40
DB 263 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAGATAC 322
QY 41 IIGLIYLYLYSARGALATNARGPHEPTRASPProArgArgGlyThrPro 58
DB 323 ATTGGCAAGAAAGAAAGCAACGACGATCTGGGATCCAGAGGGAGCAACCA 376

RESULT 2
US-09-864-761-15438
/ Sequence 15438, Application US/09864761
/ Patent No. US20020048763A1
/ GENERAL INFORMATION:
/ APPLICANT: Penn, Sharon G.
/ APPLICANT: Rank, David R.
/ APPLICANT: Hanzel, David K.
/ APPLICANT: Chen, Wensheng
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
```

```
/ TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
/ FILE REFERENCE: Aeomica-X-1
/ CURRENT APPLICATION NUMBER: US/09/864,761
/ CURRENT FILING DATE: 2001-05-23
/ PRIOR APPLICATION NUMBER: US 60/180,312
/ PRIOR FILING DATE: 2000-02-04
/ PRIOR APPLICATION NUMBER: US 60/207,456
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: US 09/632,366
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: GB 24263,6
/ PRIOR FILING DATE: 2000-10-04
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 2000-09-27
/ PRIOR APPLICATION NUMBER: PCT/US01/00666
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00667
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00664
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00669
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00665
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00668
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00663
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00662
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00661
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00670
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: US 09/774,203
/ PRIOR FILING DATE: 2001-01-29
/ NUMBER OF SEQ ID NOS: 49117
/ SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 15438
/ LENGTH: 477
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AF159056.1
/ OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.4
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.3
US-09-864-761-15438

Alignment Scores:
Pred. No.: 5,21e-36 Length: 477
Score: 307.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-031-158-14 (1-58) x US-09-864-761-15438 (1-477)

QY 1 MetGlnMetPheProPsePProLeuPhePheLeuGlnLeuLeuYsgInSerSer 20
DB 212 ATGCAAGATGTTTCCCAAGCCCACTATTTTCTTCAATTGCTGAAACAAAGCTCC 271
QY 21 ArgArgLeuGlnIuIstIrhPheValPheLeuArgAsnPheserLeuMetLeuLeuArgTyr 40
DB 272 AGAAGGCTGGAACATACCTTGTCTTCTTGAGAAATTTTCCCTGATGTTATTAAGATAC 331
QY 41 IIGLIYLYLYSARGALATNARGPHEPTRASPProArgArgGlyThrPro 58
```

```

P R I O R F I L I N G D A T E : 2001-08-22
P R I O R A P P L I C A T I O N N U M B E R : 60/325,020
P R I O R F I L I N G D A T E : 2001-09-25
P R I O R A P P L I C A T I O N N U M B E R : 60/341,746
P R I O R F I L I N G D A T E : 2001-12-12
P R I O R A P P L I C A T I O N N U M B E R : 60/362,158
P R I O R F I L I N G D A T E : 2002-03-05
N U M B E R O F S E Q I D N O S : 455
S O F T W A R E : PaSeQ for Windows Version 4.0
S E Q I D N O 404
L E N G T H : 1027
T Y P E : DNA
O R G A N I S M : Homo sapiens
U S -10-205-823-404

Alignment Scores:
Pred. No.:      1.5e-35      Length:      1027
Score:          307.00       Matches:     58
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches:    0
Query Match:     100.00%     Indels:      0
DB:              15         Gaps:        0

US-10-031-158-14 (1-58) x US-10-205-823-404 (1-1027)

Qy      1 MetGlnMetPheProProSerProLeuPhePhePheLeuGlnLeuLeuGlnSerSer 20
Db      74 ATGCACATGTTCCTCCCAAGCACCACACTATTTCCTTCCTCAATGGCGAAACAAGCTCC 133

Qy      21 ArgArgLeuGluHisThrPheValPheLeuArgAenPheSerLeuMetLeuLeuArgTyr 40
Db      134 AGAGAGCTGAACACATRCCTTTGTCTTCTTGAGAAATTTTTCCCTCATGTTTATAAGATAC 193

Qy      41 lIeGlYLyAlYaRgArGaLaThRxrGpHeTrPAepPrOAArgGlyThrPro 58
Db      194 ATTGGCAGAAGAAAAGAGCAACAGACATTCTGGATCCAGAGGGGGAACACCA 247

RESULT 5
US-10-643-795A-23
; Sequence 23, Application US/10643795A
; Publication No. US20040241703A1
GENERAL INFORMATION:
APPLICANT: FREDERIC J. DESAUVAGE
APPLICANT: GRETCHEN FRANTZ
APPLICANT: KENNETH J. HILLAN
APPLICANT: PAUL POLAKIS
APPLICANT: ANDREW POLSON
APPLICANT: VICTORIA SMITH
APPLICANT: SUSAN D. SPENCER
APPLICANT: THOMAS D. WU
APPLICANT: ZEMIN ZHANG
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
TITLE OF INVENTION: TREATMENT OF TUMOR
FILE REFERENCE: P5026R1-US
CURRENT APPLICATION NUMBER: US/10/643,795A
CURRENT FILING DATE: 2003-08-19
PRIOR APPLICATION NUMBER: US 60/404,809
PRIOR FILING DATE: 2002-08-19
PRIOR APPLICATION NUMBER: US 60/405,645
PRIOR FILING DATE: 2002-08-21
PRIOR APPLICATION NUMBER: US 60/413,192
PRIOR FILING DATE: 2002-09-23
PRIOR APPLICATION NUMBER: US 60/419,008
PRIOR FILING DATE: 2002-10-15
PRIOR APPLICATION NUMBER: US 60/426,847
PRIOR FILING DATE: 2002-11-15
PRIOR APPLICATION NUMBER: US 60/484,959
PRIOR FILING DATE: 2003-07-02
NUMBER OF SEQ ID NOS: 158
SEQ ID NO 23
LENGTH: 1395
TYPE: DNA
ORGANISM: Homo sapien

```

US-10-643-795A-23

Alignment Scores:  
Pred. No.: 2,296-35 Length: 1395  
Score: 307.00 Matches: 58  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 18 Gaps: 0

US-10-031-158-14 (1-58) x US-10-643-795A-23 (1-1395)

QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYgInSerSer 20  
DB 449 ATGCAGATGTTTCCCCCAAGCCACGATTTTCTTCTTCAATGCTGAAACAAAGCTCC 508  
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAnPheSerLeuLeuLeuArgTyr 40  
DB 509 AGAAGGCTGGAAACATACCTTTGCTTCTTGAGAAATTTTCCCTGATTTAAGATAC 568  
QY 41 IlegLYLysLYsArgArgAlaThrArgPheTTPAappProArgArgGLYThrPro 58  
DB 569 ATTGGCAAGAAAGAGAGCAACAGATTCTGGATCCAGAGGGGAAACACCA 622

RESULT 6

US-09-957-708-16 Application US/09957708  
Sequence 16, Application US/09957708  
Publication No. US20030031678A1  
GENERAL INFORMATION:  
APPLICANT: Sun, Yongming  
APPLICANT: Recipon, Heive  
APPLICANT: Cafetkey, Robert  
APPLICANT: Ali, Shujah  
TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific  
FILE REFERENCE: DEX-0239  
CURRENT APPLICATION NUMBER: US/09/957,708  
CURRENT FILING DATE: 2001-09-19  
PRIOR APPLICATION NUMBER: 60/233,746  
PRIOR FILING DATE: 2000-09-19  
NUMBER OF SEQ ID NOS: 40  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 16  
LENGTH: 1418  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-957-708-16

Alignment Scores:

Pred. No.: 2,356-35 Length: 1418  
Score: 307.00 Matches: 58  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 10 Gaps: 0

US-10-031-158-14 (1-58) x US-09-957-708-16 (1-1418)

QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYgInSerSer 20  
DB 449 ATGCAGATGTTTCCCCCAAGCCACGATTTTCTTCTTCAATGCTGAAACAAAGCTCC 508  
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAnPheSerLeuLeuLeuArgTyr 40  
DB 509 AGAAGGCTGGAAACATACCTTTGCTTCTTGAGAAATTTTCCCTGATTTAAGATAC 568  
QY 41 IlegLYLysLYsArgArgAlaThrArgPheTTPAappProArgArgGLYThrPro 58  
DB 569 ATTGGCAAGAAAGAGAGCAACAGATTCTGGATCCAGAGGGGAAACACCA 622

RESULT 7

US-09-864-761-31943  
Sequence 31943, Application US/09864761

Patent No. US20020048763A1

GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wenheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aecmica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 31943  
LENGTH: 316  
TYPE: DNA  
ORGANISM: Homo sapiens

FEATURES:

OTHER INFORMATION: MAP TO AF159056.1  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.2  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1  
OTHER INFORMATION: SWISSPROT HIT: P03986, EVALUE 7.00e-58  
OTHER INFORMATION: NT HIT: M44996.1, EVALUE 0.00e+00  
OTHER INFORMATION: EST\_HUMAN HIT: BF679123.1, EVALUE 0.00e+00  
US-09-864-761-31943

Alignment Scores:

Pred. No.: 1,646-35 Length: 316  
Score: 302.00 Matches: 57  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.37% Indels: 0  
DB: 9 Gaps: 0

US-10-031-158-14 (1-58) x US-09-864-761-31943 (1-316)

```
QY 2 GlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYsgInSerSerArg 21
| | | | |
Db 2 CAGAGTTCCTCCCAAGCCACCTATTTCTTCTTCAATTCGTGAACAAAGCTCCGA 61
QY 22 ArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyrIle 41
| | | | |
Db 62 AGGCTGGAAACATACCTTCTTCTTCTTGAAGAAATTTTCCCTGATGTTATTAAGATACATT 121
QY 42 GLyLyLyLyArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58
| | | | |
Db 122 GGCAGAGAAAGAGAGACCAACAGATTCTGGATCCAGAGGAGGAAACACCA 172

RESULT 8
US-09-954-456-317
; Sequence 317: Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
; FILE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 317
; LENGTH: 1421
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-317

Alignment Scores:
Pred. No.: 2,056-33 Length: 1421
Score: 294.00 Matches: 55
Percent Similarity: 98.28% Conservative: 2
Best Local Similarity: 94.83% Mismatches: 1
Query Match: 95.77% Indels: 0
DB: 9 Gaps: 0

US-10-031-158-14 (1-58) x US-09-954-456-317 (1-1421)
QY 1 MetGlnMetPheProSerProLeuPhePhePheLeuGlnLeuLeuYsgInSerSer 20
| | | | |
Db 423 ATGCAAGATGTTCTCCCAAGCCACCTATTTCTTCTTCAATTCGTGAACAAAGCTCC 482
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40
| | | | |
Db 483 AGAAGGCTGGAAACATACCTTGTCTTCTTGAAGAAATTTTCCAGATATTATTAAATAC 542
QY 41 ILeGlyLyLyArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58
| | | | |
Db 543 ATTGGCAAGAAAGAGAGACCAACAGATTCTGGATCCAGAGGAGGAAACACCA 596
```

```
RESULT 9
US-10-101-510-492
; Sequence 492: Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; APPLICANT: WAN, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117,0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 492
; LENGTH: 1799
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-492

Alignment Scores:
Pred. No.: 2,846-33 Length: 1799
Score: 294.00 Matches: 58
Percent Similarity: 98.31% Conservative: 0
Best Local Similarity: 95.77% Mismatches: 1
Query Match: 95.77% Indels: 0
DB: 15 Gaps: 0

US-10-031-158-14 (1-58) x US-10-101-510-492 (1-1799)
QY 1 MetGlnMetPheProSerProLeuPhePhePheLeuGlnLeuLeuYsgInSerSer 20
| | | | |
Db 845 ATGCAAGATGTTCTCCCAAGCCACCTATTTCTTCTTCAATTCGTGAACAAAGCTCC 904
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40
| | | | |
Db 905 AGAAGGCTGGAAACATACCTTGTCTTCTTGAAGAAATTTTCCCTGATGTTATTAAGATAC 964
QY 41 ILeGlyLyLyArgArgAlaThrArgPheTyrPheProArgArgGlyThrPro 58
| | | | |
Db 965 ATTGGCAAGAAAGAGAGACCAACAGATTCTGGATCCAGAGGAGGAAACACCA 1019

RESULT 10
US-09-960-706-676
; Sequence 676: Application US/09960706
; Publication No. US20030134280A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William B.
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplas
; FILE REFERENCE: 44921-5029-01US
; CURRENT APPLICATION NUMBER: US/09/960,706
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 60/223,323
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 09/873,319
; PRIOR FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 676
; LENGTH: 1586
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134280A1 M30894
US-09-960-706-676

Alignment Scores:
Pred. No.: 7,436-32 Length: 1586
Score: 284.00 Matches: 53
Percent Similarity: 96.55% Conservative: 3
Best Local Similarity: 91.38% Mismatches: 2
```

```
Query Match: 92.51% Indels: 0
DB: 10 Gaps: 0
US-10-031-158-14 (1-58) x US-09-960-706-676 (1-1586)
QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYsgInSerSer 20
DB 553 ATGCAGATGTTTCCCCCAAGCCACATATTTTCTTCCTTCGATTCGTAACAACTCC 612
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40
DB 613 AGAAGGCTGGAAACATCTTTGCTCTTGAGAAATTTTCCAGATATTATTAAATAC 672
QY 41 IleGlyLysLysArgArgAlaThrArgPheTyrAspProArgArgGlyThrPro 58
DB 673 ATTGGCAAGAAAGAAAGACACACGATTCGGGATCCCGAGGGGAAACACCA 726

RESULT 11
US-09-873-319-424
; Sequence 424, Application US/09873319A
; Publication No. US20030134324A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William F.
; APPLICANT: Kulkaril, Prakash
; APPLICANT: Getzenberg, Robert H.
; APPLICANT: Mago, Iwao
; APPLICANT: Yamamoto, Jun
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic
; TITLE OF INVENTION: Hyperplasia Using Gene Expression Profiles
; FILE REFERENCE: 44921-5029-US
; CURRENT APPLICATION NUMBER: US/09/873,319A
; EARLIER FILING DATE: 2001-06-05
; EARLIER APPLICATION NUMBER: US 60/223,323
; EARLIER FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 755
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 424
; LENGTH: 1586
; TYPE: DNA
; ORGANISM: Homo sapiens
; OTHER INFORMATION: Genbank Accession No. US20030134324A1 M30894
US-09-873-319-424

Alignment Scores:
Pred. No.: 7,43e-32 Length: 1586
Score: 284.00 Matches: 53
Percent Similarity: 96.55% Conservative: 3
Best Local Similarity: 91.38% Mismatches: 2
Query Match: 92.51% Indels: 0
Gaps: 0

US-10-031-158-14 (1-58) x US-09-873-319-424 (1-1586)
QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYsgInSerSer 20
DB 553 ATGCAGATGTTTCCCCCAAGCCACATATTTTCTTCCTTCGATTCGTAACAACTCC 612
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40
DB 613 AGAAGGCTGGAAACATCTTTGCTCTTGAGAAATTTTCCAGATATTATTAAAGTAC 672
QY 41 IleGlyLysLysArgArgAlaThrArgPheTyrAspProArgArgGlyThrPro 58
DB 673 ATTGGCAAGAAAGAAAGACACACGATTCGGGATCCCGAGGGGAAACACCA 726

RESULT 12
US-09-925-300-58
; Sequence 58, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
```

```
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 58
; LENGTH: 1155
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (135)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (432)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (443)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-58

Alignment Scores:
Pred. No.: 3.19e-26 Length: 1155
Score: 245.00 Matches: 51
Percent Similarity: 91.07% Conservative: 0
Best Local Similarity: 91.07% Mismatches: 5
Query Match: 79.80% Indels: 1
Gaps: 0

US-10-031-158-14 (1-58) x US-09-925-300-58 (1-1155)
QY 1 MetGlnMetPheProSerProLeuPhePheLeuGlnLeuLeuYsgInSerSer 20
DB 175 ATGCAGATGTTTCCCCCAAGCCACATATTTCTTCCTTCGATTCGTAACAACTCC 233
QY 21 ArgArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeuArgTyr 40
DB 234 AGAAGGCTGGAAACATCTTTGCTCTTGAGAAATTTTCCAGATATTATTAAATAC 293
QY 41 IleGlyLysLysArgArgAlaThrArgPheTyrAspProArgArgGly 56
DB 294 ATTGGSAAGAAAGAAAGACACACGATTCGGGATCCCGAGGGGG 341

RESULT 13
US-10-437-963-35632/C
; Sequence 35632, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; PRIOR FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 35632
; LENGTH: 1813
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_39533C.1
```

US-10-437-963-35632

## Alignment Scores:

Pred. No.:	1.43	Length:	1813
Score:	75.00	Matches:	24
Percent Similarity:	40.85%	Conservative:	5
Best Local Similarity:	33.80%	Mismatches:	22
Query Match:	24.43%	Indels:	20
DB:	17	Gaps:	3

US-10-031-158-14 (1-58) x US-10-437-963-35632 (1-1813)

QY 5 Propio-----SerProLeuPhePheLeuGlnLeuLeuys 17

DB 1740 CCGCCACCGCTATATACTGTCGCTTACGCCCTTATTTTCTCTCTCCATCTCT 1681

QY 18 GlnSer-----ArgArgLeuGlnHis 25

DB 1680 TCGTCTCTCTCTCTCTCTCTCTCTCCACACTGCACACTCTGTGATCT 1621

QY 26 ThrPheValPheLeuArgAnPheSerLeuMetLeuArgTyrIleGlyIysArg 45

DB 1620 CTCCTCTCTCTCTGTATAT--TGCCTCTCTCTCTCTCGATCTCGCGCATCA 1564

QY 46 ArgAlaThrArgPheThrPheProArgArgGly 56

DB 1563 TGGCCATGGAGAGTGAACGCCAAGAGAGG 1531

## RESULT 14

US-10-027-632-136731/c

; Sequence 136731, Application US/10027632

; Publication No. US20020198371A1

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632

; PRIOR FILING DATE: 2002-04-30

; PRIOR APPLICATION NUMBER: US 60/218,006

; PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 60/198,676

; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: US 60/193,483

; PRIOR FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23

; PRIOR APPLICATION NUMBER: US 60/156,358

; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002

; NUMBER OF SEQ ID NOS: 325720

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO: 136731

; LENGTH: 618

; TYPE: DNA

; ORGANISM: Human

US-10-027-632-136731

## Alignment Scores:

Pred. No.:	0.54	Length:	618
Score:	73.50	Matches:	22
Percent Similarity:	51.72%	Conservative:	8
Best Local Similarity:	37.93%	Mismatches:	25
Query Match:	23.94%	Indels:	3
DB:	13	Gaps:	1

US-10-031-158-14 (1-58) x US-10-027-632-136731 (1-618)

QY 3 MetPheProSerProLeuPhePheLeuGlnLeuLeuysGlnSerArgArg 22

:::||||| ||| ||||| ||:::

DB 195 CTATTCACCTCCCAACTGCTTTTTTTT---GAGACARAGTCTGCTGTCAACAG 139

QY 23 LeuGlnHisThrPheValPheLeuArgAnPheSerLeuMetLeuArgTyrIleGly 42

DB 138 CTGGAGCATAGATGATGATCTTGCTCACTGCACACTCTGCTCTGAGTTCAAGCAAT 79

QY 43 LysIysArgArgAlaThrArgPhe--TTPAsPProArgArgGlyThrPro 58

DB 78 TCTCTGCTCAGCCTCCCAAGTACTGAGACTACAGAGGCGCACCA 29

## RESULT 15

US-10-027-632-136731/c

; Sequence 136731, Application US/10027632

; Publication No. US200204075A9

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632

; PRIOR FILING DATE: 2002-04-30

; PRIOR APPLICATION NUMBER: US 60/218,006

; PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 60/198,676

; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: US 60/193,483

; PRIOR FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23

; PRIOR APPLICATION NUMBER: US 60/156,358

; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002

; NUMBER OF SEQ ID NOS: 325720

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO: 136731

; LENGTH: 618

; TYPE: DNA

; ORGANISM: Human

US-10-027-632-136731

## Alignment Scores:

Pred. No.:	0.54	Length:	618
Score:	73.50	Matches:	22
Percent Similarity:	51.72%	Conservative:	8
Best Local Similarity:	37.93%	Mismatches:	25
Query Match:	23.94%	Indels:	3
DB:	15	Gaps:	1

US-10-031-158-14 (1-58) x US-10-027-632-136731 (1-618)

QY 3 MetPheProSerProLeuPhePheLeuGlnLeuLeuysGlnSerArgArg 22

DB 195 CTATTCACCTCCCAACTGCTTTTTTTT---GAGACARAGTCTGCTGTCAACAG 139

QY 23 LeuGlnHisThrPheValPheLeuArgAnPheSerLeuMetLeuArgTyrIleGly 42

DB 138 CTGGAGCATAGATGATGATCTTGCTCACTGCACACTCTGCTCTGAGTTCAAGCAAT 79

QY 43 LysIysArgArgAlaThrArgPhe--TTPAsPProArgArgGlyThrPro 58

DB 78 TCTCTGCTCAGCCTCCCAAGTACTGAGACTACAGAGGCGCACCA 29

## RESULT 16

US-10-027-632-105504/c

; Sequence 105504, Application US/10027632

; Publication No. US20020198371A1

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

; FILE REFERENCE: Polymorphisms in the Human Genome

```
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 105504
LENGTH: 618
TYPE: DNA
ORGANISM: Human
US-10-027-632-105504

Alignment Scores:
Pred. No.: 0.761 Length: 618
Score: 72.50 Matches: 22
Percent Similarity: 51.72% Conservative: 8
Best Local Similarity: 37.93% Mismatches: 25
Query Match: 23.62% Indels: 3
DB: 13 Gaps: 1

US-10-031-158-14 (1-58) x US-10-027-632-105504 (1-618)

QY 3 MetPheProPserProLeuPhePheLeuGlnLeuLeuYsgInserSerArgArg 22
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 195 CTAATCACTCCCAAACTGCTTTT---GAGACAAAGCTGCTCTGTACACAGG 139
QY 23 LeuGlnHieThrPheValPheLeuArgAenPheSerLeuMetLeuLeuArgTyrlleGly 42
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 138 CTGAGACATAGTGAATGATCTTGCTGCTGCAACTGCTGCTCTGTTCAAGCAAT 79
QY 43 LysLysArgArgAlaThrArgPhe--TTPApProArgArgGlyThrPro 58
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 78 TCTCCTGCTCAGCTCCCAAGTAGCTGGAGACTACAGAGGGGCCACCA 29

RESULT 17
US-10-027-632-105505/c
; Sequence 105505, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; CURRENT APPLICATION NUMBER: US/10/027,632
; SEQ ID NO 105504
; LENGTH: 618
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-105504

Alignment Scores:
Pred. No.: 0.761 Length: 618
Score: 72.50 Matches: 22
Percent Similarity: 51.72% Conservative: 8
Best Local Similarity: 37.93% Mismatches: 25
Query Match: 23.62% Indels: 3
DB: 15 Gaps: 1

US-10-031-158-14 (1-58) x US-10-027-632-105504 (1-618)

QY 3 MetPheProPserProLeuPhePheLeuGlnLeuLeuYsgInserSerArgArg 22
```

```
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 105505
LENGTH: 618
TYPE: DNA
ORGANISM: Human
US-10-027-632-105505

Alignment Scores:
Pred. No.: 0.761 Length: 618
Score: 72.50 Matches: 22
Percent Similarity: 51.72% Conservative: 8
Best Local Similarity: 37.93% Mismatches: 25
Query Match: 23.62% Indels: 3
DB: 13 Gaps: 1

US-10-031-158-14 (1-58) x US-10-027-632-105505 (1-618)

QY 3 MetPheProPserProLeuPhePheLeuGlnLeuLeuYsgInserSerArgArg 22
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 195 CTAATCACTCCCAAACTGCTTTT---GAGACAAAGCTGCTCTGTACACAGG 139
QY 23 LeuGlnHieThrPheValPheLeuArgAenPheSerLeuMetLeuLeuArgTyrlleGly 42
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 138 CTGAGACATAGTGAATGATCTTGCTGCTGCAACTGCTGCTCTGTTCAAGCAAT 79
QY 43 LysLysArgArgAlaThrArgPhe--TTPApProArgArgGlyThrPro 58
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 78 TCTCCTGCTCAGCTCCCAAGTAGCTGGAGACTACAGAGGGGCCACCA 29

RESULT 18
US-10-027-632-105504/c
; Sequence 105504, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; CURRENT APPLICATION NUMBER: US/10/027,632
; SEQ ID NO 105504
; LENGTH: 618
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-105504

Alignment Scores:
Pred. No.: 0.761 Length: 618
Score: 72.50 Matches: 22
Percent Similarity: 51.72% Conservative: 8
Best Local Similarity: 37.93% Mismatches: 25
Query Match: 23.62% Indels: 3
DB: 15 Gaps: 1

US-10-031-158-14 (1-58) x US-10-027-632-105504 (1-618)

QY 3 MetPheProPserProLeuPhePheLeuGlnLeuLeuYsgInserSerArgArg 22
```

```

Db      195 CTATTCACCTCCCAAACTGCTTTTTTTT---GAGACAAGCTCCTCGTGTACACAGG 139
Oy      23 LeuglunhiStrPhyeValPheleuAgaenPheSerMetLeuLeuArgTyrllegly 42
Db      138 CTGGAGCATGTAGTATGTGATCTTGCTGACTGCACCTGCTGCTGCTGATTCAAGAAT 79
Oy      43 LyselysaRgAlathrArqPhe--TrpAspProArqArgGlYthrPro 58
Db      78 TCTCTGCCCTCACGCTTCCCACAGTAGCTGGAGCTACAGAGAGGCCGCCACCA 29

RESULT 19
US-10-027-632-105505/c
; Sequence 105505, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ. ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 105505
; LENGTH: 618
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-105505

Alignment Scores:
Pred. No.:
Score: 0.761 Length: 618
Percent Similarity: 72.52% Matches: 22
Best Local Similarity: 37.93% Mismatches: 25
Query Match: 23.62% Indels: 3
DB: 15 Gaps: 1

US-10-031-158-14 (1-58) x US-10-027-632-105505 (1-618)
Oy      3 MetPheProSerProIeuPhePheLuglinLeuLySGlnSerSerArgarg 22
Db      195 CTATTCACCTCCCAAACTGCTTTTTTTT---GAGACAAGCTCCTCGTGTACACAGG 139
Oy      23 LeuglunhiStrPhyeValPheleuAgaenPheSerMetLeuLeuArgTyrllegly 42
Db      138 CTGGAGCATGTAGTATGTGATCTTGCTGACTGCACCTGCTGCTGCTGATTCAAGAAT 79
Oy      43 LyselysaRgAlathrArqPhe--TrpAspProArqArgGlYthrPro 58
Db      78 TCTCTGCCCTCACGCTTCCCACAGTAGCTGGAGCTACAGAGAGGCCGCCACCA 29

RESULT 20
US-09-764-991-10230/c
; Sequence 10230, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

```

```

; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10230
; LENGTH: 27118
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-10230

Alignment Scores:
Pred. No.:          141          Length: 27118
Score:              72.50        Matches: 19
Percent Similarity: 64.86%      Conservative: 5
Best Local Similarity: 51.35%   Mismatches: 10
Query Match:        23.62%      Indels: 3
DB:                 10         Gaps: 1

US-10-031-158-14 (1-58) x US-09-764-891-10230 (1-27118)

Qy      5  ProProSerProLeuphePhePheLeuGlnLeuLeuGlnSer-----SerArg 21
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      15017 CCACGACGCCCGGCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTCTGCTGCGCC 14958
Qy      22  ArgLeuGlnHisThrPheValPheLeuArgAsnPheSerLeuMetLeuLeu 38
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      14957 AGGCTGAGAGTCAGATGGTGTGATCTTGCTGCTCACTGCACCTCTGTCTCTTA 14907

RESULT 1
US-10-205-428-1018/c
; Sequence 1018, Application US/10205428
; Publication No. US20030108907A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA117C1
; CURRENT APPLICATION NUMBER: US/10/205,428
; CURRENT FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: 09/764,892
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1019
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1018
; LENGTH: 27118
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-205-428-1018

Alignment Scores:
Pred. No.:          141          Length: 27118
Score:              72.50        Matches: 19
Percent Similarity: 64.86%      Conservative: 5
Query Match:        23.62%      Indels: 3
DB:                 10         Gaps: 1

```

[illegible]

```
Alignment Scores:
Pred. No.: 1 51 Length: 618
Score: 70.50 Matches: 22
Percent Similarity: 50.00% Conservative: 7
Best Local Similarity: 37.93% Mismatches: 26
Query Match: 22.96% Indels: 3
DB: 15 Gaps: 1

US-10-031-158-14 (1-58) x US-10-027-632-136730 (1-618)

QY 3 MetPheProSerProLeuPhePheLeuGlnLeuLysGlnSerSerArgArg 22
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 195 CTAATCAGTCCCAACCTGCTTTT---GAGACAAAGCTCTGTCACACAG 139

QY 23 LeuGlnHisThrPheValPheLeuArgAnPheSerLeuMetLeuLeuArgTyrIleGly 42
   |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 138 CTGAGCAGTATGATGATCTTGCTCACTGCAACCTGCTGCTGCTGCTCAAGCANT 79

QY 43 LysLysArgArgAlaThrArgPhe--TyrAspProArgArgGlyThrPro 58
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 78 TCTCTGCTCAGCTCTCCAGTACTGAGTACTGAGAGGCGCCACCA 29

RESULT 25
US-10-424-599-127254/c
; Sequence 127254, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 127254
; LENGTH: 353
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_85916C.1
US-10-424-599-127254

Alignment Scores:
Pred. No.: 0.985 Length: 353
Score: 69.50 Matches: 23
Percent Similarity: 50.88% Conservative: 6
Best Local Similarity: 40.35% Mismatches: 21
Query Match: 22.64% Indels: 7
DB: 16 Gaps: 3

US-10-031-158-14 (1-58) x US-10-424-599-127254 (1-353)

QY 5 ProProSerProLeuPhePheLeuGlnLeuLysGlnSer--SerArgArgLeu 23
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 285 CCCCCCTCCCCCTATTTTTCCTGCGGATGCGCCAGTAATGTATACCCCCGCTG 226

QY 24 GlnHisThrPheValPheLeuArgAnPheSer-----LeuMetLeuLeuArg 39
   |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 225 TGGGAAGACGATGATTTTGTGCGGGGTTAGTGCCAAATATATAGTTAATGTGTGCATC 166

QY 40 TyrIleGlyLysLysArgArgAlaThrArgPheThrAspProArgArgGly 56
   |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 165 GAAATTGGGGTAGTAGGGAAGAGCCCTTC-----CCCAAAAAGGA 121

RESULT 26
US-10-282-122A-22460/c
; Sequence 22460, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haebelbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patencin version 3.1
; SEQ ID NO 22460
; LENGTH: 2145
; TYPE: DNA
; ORGANISM: Helicobacter pylori
US-10-282-122A-22460

Alignment Scores:
Pred. No.: 11.9 Length: 2145
Score: 69.50 Matches: 25
Percent Similarity: 51.61% Conservative: 7
Best Local Similarity: 40.32% Mismatches: 17
Query Match: 22.64% Indels: 14
DB: 16 Gaps: 3

US-10-031-158-14 (1-58) x US-10-282-122A-22460 (1-2145)

QY 1 MetGlnMetPheProProSerPro-----Leu 9
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 2079 ATACAAAGTTCCTCCCTGACCTGATCTTCTAAACCATATTCCTAGCCTT 2020

QY 10 PhePhePheLeuGlnLeuLysGlnSerSerArgArgLeuGlnHisThrPheValPhe 29
   |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 2019 GATATTTCTTGCA---TTCAAAAGGATGAGCCAGTGGACTTCTATTTCTGCTTT 1963

QY 30 LeuArgAnPheSerLeuMetLeuLeuArgTyrIleGlyLysLysArgArgAlaThrArg 49
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1962 TCCAGCT-TTT---CTAGTCTTAATATCGTTTCAAAATCTCAGCGCTCTCTT 1907

QY 50 PheTyr 51
   |||||
Db 1906 TTTTGG 1901

RESULT 27
US-09-881-752A-321/c
```

```
/ Sequence 321, Application US/09881752A
/ Patent No. US20020115078A1
/ GENERAL INFORMATION:
/ APPLICANT: Kleantous, Harold
/ APPLICANT: Al-Garawi, Amal
/ APPLICANT: Miller, Charles
/ APPLICANT: Tomb, Jean-Francois
/ APPLICANT: Oomen, Raymond P.
/ TITLE OF INVENTION: Identification of Polynucleotides
/ TITLE OF INVENTION: Encoding No. US20020115078A1 Helicobacter Polypeptides in the
/ TITLE OF INVENTION: Genome
/ FILE REFERENCE: 06132/041002
/ CURRENT APPLICATION NUMBER: US/09/881,752A
/ PRIOR FILING DATE: 2001-06-15
/ PRIOR APPLICATION NUMBER: US 08/833,457
/ PRIOR FILING DATE: 1997-04-01
/ NUMBER OF SEQ ID NOS: 370
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 321
/ LENGTH: 2241
/ TYPE: DNA
/ ORGANISM: Helicobacter pylori
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (52)...(2193)
/ US-09-881-752A-321

Alignment Scores:
Pred. No.: 12.7 Length: 2241
Score: 69.50 Matches: 25
Percent Similarity: 51.61% Conservative: 17
Best Local Similarity: 40.32% Mismatches: 7
Query Match: 22.64% Indels: 14
DB: 9 Gaps: 3

US-10-031-158-14 (1-58) x US-09-881-752A-321 (1-2241)

Qy 1 MetGlnMetPheProPheProSerPro-----Leu 9
   |||||
Db 2130 ATACAAATGTTCCCTTACCTGCTTCTTAAAGCATATTCTAGCCTT 2071
   |||||

Qy 10 PhePhePheLeuGlnLeuLeuLysGlnSerSerArgArgLeuGlnHisThrPheValPhe 29
   |||||
Db 2070 GATATTTTCTTGCAA--TTCAAAGGCGATGAGGCGATGTGACCTTCTTCTTCTT 2014
   |||||

Qy 30 LeuArgAsnPheSerLeuMetLeuLeuArgTyrIleGlyLysLysArgAlaThrArg 49
   |||||
Db 2013 TCCAGGCT-TTT--CTAGTTCTAAAAATCGTTTCAAAATCTCAGCGCTCTTCTT 1958
   |||||

Qy 50 PheTyr 51
   |||||
Db 1957 TTTTGG 1952

RESULT 28
US-10-027-632-203503
/ Sequence 203503, Application US/10027632
/ Publication No. US2002019837A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ TITLE OF INVENTION: Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ PRIOR FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/199,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
```

```
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 203503
/ LENGTH: 585
/ TYPE: DNA
/ ORGANISM: Human
/ US-10-027-632-203503

Alignment Scores:
Pred. No.: 2.79 Length: 585
Score: 68.50 Matches: 18
Percent Similarity: 49.09% Conservative: 9
Best Local Similarity: 32.73% Mismatches: 19
Query Match: 22.31% Indels: 9
DB: 13 Gaps: 2

US-10-031-158-14 (1-58) x US-10-027-632-203503 (1-585)

Qy 5 ProProSerProLeuPhePhePheLeuGlnLeuLysGlnSerSerArgArgLeuGln 24
   |||||
Db 361 CCTCCCTTCCCTCCCTCCACTTTC-----CAATCTCTTCTCTTATCC 405
   |||||

Qy 25 HisThrPheValPheLeuArgAsnPheSerLeuMetLeu-----ArgTyr 40
   |||||
Db 406 CTAACTTATTTATTTATTTTGGATTCATCTTATATCTGTACACACTTGAACATTAT 465
   |||||

Qy 41 IleGlyLysLysArgArgAlaThrArgPheTyrPheProArgArg 55
   |||||
Db 466 GTAGGAAAAAGAGCTGCGATTTCCTTATGTAAACCAAGAAA 510
   |||||

RESULT 29
US-10-027-632-203503
/ Sequence 203503, Application US/10027632
/ Publication No. US20030204075A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ TITLE OF INVENTION: Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ PRIOR FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 203503
/ LENGTH: 585
/ TYPE: DNA
/ ORGANISM: Human
/ US-10-027-632-203503

Alignment Scores:
Pred. No.: 2.79 Length: 585
Score: 68.50 Matches: 18
Percent Similarity: 49.09% Conservative: 9
Best Local Similarity: 32.73% Mismatches: 19
```

```

Query Match:      22.31%      Indels:      9
DB:               15          Gaps:         2

US-10-031-158-14 (1-58) x US-10-027-632-203503 (1-585)

QY      5 ProPterSerProLeuPhePhePheLeuGlnLeuLeuGlnSerSerArgArgLeuGln 24
      |||||
DB      361 CCGCCCTTCCCTCCCTCCACTTTTC-----CAATCTCTCTCTTTATATCC 405
      |||||

QY      25 HistH-PheValPheLeuArgPhePheSerLeuMetLeu-----ArgTyr 40
      |||||
DB      406 CTACACTTATATATTTTGTGATTCACCTTATCTGTACACACTTGACCAATAT 465
      |||||

QY      41 IleGlyLeuArgArgArgArgArgPheThrArgPheThrArgArgArg 55
      |||||
DB      466 GTAGGAAAGAGGCTGCGCATTTTCTTATTGTACCAAGAA 510
      |||||

RESULT 30
US-10-363-345A-15079      Application US/10363345A
; Sequence 15079, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE REFERENCE: E01/1227
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 15079
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: Cpg-island No: 15079
US-10-363-345A-15079

Alignment Scores:
Pred. No.:      3.48      Length:      535
Score:          67.50      Matches:      18
Percent Similarity: 55.56%      Conservative: 7
Best Local Similarity: 40.00%      Mismatches: 15
Query Match:    21.99%      Indels:      5
DB:             18          Gaps:         2

US-10-031-158-14 (1-58) x US-10-363-345A-15079 (1-535)

QY      12 PheLeuGlnLeuLeuGlnSerSerArgArgLeuGlnHistHrPheValPheLeuArg 31
      |||||
DB      82 TTTTTCATCGTGAAGCG-----CGCGTTTAGTTTCGTTTTTTTTTTGTCGG 135
      |||||

QY      32 AsnPheSerLeuMetLeuLeuArgTyrIleGlyLeuArgArgArgArgArgPheThr 51
      |||||
DB      136 GTTTGCGGTATATTTTATTTGTTGGATCGCGTTTTCGGGTTTCGTAAGTT--- 192
      |||||

QY      52 AspProArgArgGly 56
      |||||
DB      193 -----CGTCGGGGA 201
      |||||

RESULT 31
US-10-363-345A-15080/c
; Sequence 15080, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE REFERENCE: E01/1227
; OTHER INFORMATION: cytosines in genomic DNA in the sequence context of 5'-CpG-3'

```

```

; CURRENT APPLICATION NUMBER: US/10/363,345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 15080
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: Cpg-island No: 15080
US-10-363-345A-15080

Alignment Scores:
Pred. No.:      3.48      Length:      535
Score:          67.50      Matches:      18
Percent Similarity: 55.56%      Conservative: 7
Best Local Similarity: 40.00%      Mismatches: 15
Query Match:    21.99%      Indels:      5
DB:             18          Gaps:         2

US-10-031-158-14 (1-58) x US-10-363-345A-15080 (1-535)

QY      12 PheLeuGlnLeuLeuGlnSerSerArgArgLeuGlnHistHrPheValPheLeuArg 31
      |||||
DB      454 TTTTTCGATCGTGAAGCG-----CGCGTTTAGTTTCGTTTTTTTTTTGTCGG 401
      |||||

QY      32 AsnPheSerLeuMetLeuLeuArgTyrIleGlyLeuArgArgArgArgArgPheThr 51
      |||||
DB      400 GTTTGCGGTATATTTTATTTGTTGGATCGCGTTTTCGGGTTTCGTAAGTT--- 344
      |||||

QY      52 AspProArgArgGly 56
      |||||
DB      343 -----CGTCGGGGA 335
      |||||

RESULT 32
US-09-962-832-256
; Sequence 256, Application US/09962832
; Patent No. US20020110821A1
; GENERAL INFORMATION:
; APPLICANT: Eder, Reinhard
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signat
; FILE REFERENCE: 689290-74
; CURRENT APPLICATION NUMBER: US/09/962,832
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,077
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,280
; NUMBER OF SEQ ID NOS: 259
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 256
; LENGTH: 942
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-832-256

Alignment Scores:
Pred. No.:      7.6      Length:      942
Score:          67.50      Matches:      17
Percent Similarity: 61.11%      Conservative: 5
Best Local Similarity: 47.22%      Mismatches: 11
Query Match:    21.99%      Indels:      3
DB:             9          Gaps:         2

US-10-031-158-14 (1-58) x US-09-962-832-256 (1-942)

QY      21 ArgArgLeuGlnHistHrPheValPheLeuArgAsnPheSerLeuMetLeu-----Arg 39
      |||||
DB      620 AACAATCTGCTGCATACCGCTATATGATGATGATTTTCCATGCCCTGTACGGGAGA 679
      |||||

QY      40 TyrIleGlyLeuArgArgArgArgArgArgPheThrArgPheThrArgProArgArg 55
      |||||
DB      400 -----CGTCGGGGA 335
      |||||

```

```

Db      680  TACGTAGGAAAGAAAAGAGAGG-----CAGTACTGGCCATGGAGAGAAA 721

RESULT 33
US-10-172-118-1338
; Sequence 1338, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernarde, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1338
; LENGTH: 1037
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_006912
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1338

Alignment Scores:
Pred. No.:      8.68      Length:      1037
Score:          67.50      Matches:      17
Percent Similarity: 61.1%  Conservative: 5
Best Local Similarity: 47.22% Mismatches:      11
Query Match:      21.99%  Indels:         3
DB:               15      Gaps:         2

US-10-031-158-14 (1-58) x US-10-172-118-1338 (1-1037)

Qy      21  ATGATGAGLEUGLWHTSHRPhenValPheLeuArgAsnPheserLeuettLeu---Arg 39
Db      620  AAGACATCTCTGTCATACCCGCTACTATATTCATGATGTTTCCATGCCCTTGATGAGAGA 679
Qy      40  TTTTleGlyLysLysArgArgAlaThrArgPheThrAspProArgArg 55
Db      680  TACGTAGGAAAGAAAAGAGAGG-----CAGTACTGGCCATGGAGAGAAA 721

RESULT 34
US-10-342-887-1338
; Sequence 1338, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernarde, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699

```

```

; SEQ ID NO 1338
; LENGTH: 1037
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-1338

Alignment Scores:
Pred. No.:      8.68           Length:       1037
Score:          67.50         Matches:        17
Percent Similarity: 61.11%     Conservative:   5
Best Local Similarity: 47.22%    Mismatches:    11
Query Match:     21.99%        Indels:         3
DB:              16            Gaps:           2

US-10-031-158-14 (1-58) x US-10-342-887-1338 (1-1037)

Qy      21 ArgArgLeuGlnHisThrPheValPheLeuAlrGsAnPheserLauMetLeuIeu---Arg 39
Db      620 AGACATCTGCTGCATACCAGCCGTACTATTGGATGATGGTTCCTCAGCCCTTGATCGGAGA 679
               |||||
Qy      40 Tyrtllgilylylvalargalathrarqghetrpapaproaaarg 55
Db      680 TACGTAGGAAGAAAGAGAG-----CAGTACTGCCATCGAGAAAA 721
               :::::|||

RESULT 35
US-09-918-624B-4
; Sequence 4, Application US/09918624B
; Publication No. US20030113720A1
; GENERAL INFORMATION:
; APPLICANT: Scheye, Xiao Min
; TITLE OF INVENTION: CDNAS EXPRESSED IN ADIPOCYTE DIFFERENTIATION
; FILE REFERENCE: PA-0033 US
; CURRENT FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: 60/222,470
; PRIORITY FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 2532
; TYPE: DNA
; ORGANISM: Homo sapLens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: incyte ID No. US20030113720A1 481536.3
US-09-918-624B-4

Alignment Scores:
Pred. No.:      29.8           Length:       2532
Score:          67.50         Matches:        17
Percent Similarity: 61.11%     Conservative:   5
Best Local Similarity: 47.22%    Mismatches:    11
Query Match:     21.99%        Indels:         3
DB:              10            Gaps:           2

US-10-031-158-14 (1-58) x US-09-918-624B-4 (1-2532)

Qy      21 ArgArgLeuGlnHisThrPheValPheLeuArGsAnPheserLauMeleuIeu---Arg 39
Db      732 AGACATCTGCTGCATACCAGCCGTACTATTGGATGATGGTTCCTCAGCCCTTGATCGGAGA 791
               |||||
Qy      40 Tyrtllgilylylvalargalathrarqghetrpapaproaaarg 55
Db      792 TACGTAGGAAGAAAGAGAG-----CAGTACTGCCATCGAGAAAA 833
               :::::|||

RESULT 36
US-09-764-891-7020
; Sequence 7020, Application US/09744891
; Publication No. US2003007780BA1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
```

```
FILE REFERENCE: PC006
CURRENT APPLICATION NUMBER: US/09/764,891
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 10231
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7020
LENGTH: 11838
TYPER: DNA
ORGANISM: Homo sapiens
US-09-764-891-7020

Alignment Scores:
Pred. No.: 251          Length: 11838
Score: 67.50           Matches: 17
Percent Similarity: 61.1% Conservative: 5
Best Local Similarity: 47.22% Mismatches: 11
Query Match: 21.99% Indels: 3
DB: 10 Gaps: 2

US-10-031-158-14 (1-58) x US-09-764-891-7020 (1-11838)
QY 21 ArgArgLeuGluHisThrPheValPheLeuArgAspPheSerLeuMetLeu---Arg 39
Db 10840 AGACATCTGCTGCATCCGCTACTATATTCATGATGTTTCCATGCCCTTGTCGGGAGA 10899
QY 40 TyrTlGlyLysLysArgAlaThrArgPheTyrAspProArgArg 55
Db 10900 TACGTAGAAAGAAAGAGAG-----CAGTACTGCACATGAGAAA 10941

RESULT 37
US-10-311-455-1165
Sequence 1165, Application US/10311455
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Cytosine Methylation of Specific Genes
FILE REFERENCE: 5013.1014
CURRENT APPLICATION NUMBER: US/10/311,455
CURRENT FILING DATE: 2002-12-16
PRIOR APPLICATION NUMBER: PCT/EP01/07537
PRIOR FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 2424
SEQ ID NO 1165
LENGTH: 13814
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
NAME/KEY: unsure
LOCATION: 4518
OTHER INFORMATION: n 1e a o r g o r c o r t
US-10-311-455-1165

Alignment Scores:
Pred. No.: 311          Length: 13814
Score: 67.50           Matches: 18
Percent Similarity: 55.56% Conservative: 7
Best Local Similarity: 40.00% Mismatches: 15
Query Match: 21.99% Indels: 5
DB: 15 Gaps: 2

US-10-031-158-14 (1-58) x US-10-311-455-1165 (1-13814)
QY 12 PheLeuGlnLeuLeuYsgInserSerArgLeuGlnHisThrPheValPheLeuArg 31
```

```
Db 5345 TTTTTCATGCTGAAGCG-----CGCCTTACGTTTGGTTTCTTCGCG 5398
QY 32 AsnPheSerLeuMetLeuLeuArgTyrTlGlyLysLysArgAlaThrArgPheTyr 51
Db 5399 GTTTCGCGTATTTTATTTTATTTGTTGGATCGCGCTTTCGGGTTGTAAGTTT--- 5455
QY 52 AspProArgArgGly 56
Db 5456 -----CGTCGGGGA 5464

RESULT 38
US-10-425-115-4936/c
Sequence 4936, Application US/10425115
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with the Development of Plants
FILE REFERENCE: 38-21(5322)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 4936
LENGTH: 314
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_104489C.1
US-10-425-115-4936

Alignment Scores:
Pred. No.: 1.98          Length: 314
Score: 67.00           Matches: 20
Percent Similarity: 45.76% Conservative: 7
Best Local Similarity: 33.90% Mismatches: 12
Query Match: 21.82% Indels: 12
DB: 18 Gaps: 3

US-10-031-158-14 (1-58) x US-10-425-115-4936 (1-314)
QY 5 ProProSerProLeuPhePheLeuGln-----LeuLeuYsgInserSerArg 21
Db 311 CCCCCACCCCGAGGTTTTCATGCCCAATGCCCTTTCGAGTTTGCCTTCCAA 252
QY 22 ArgLeuGlnHisThrPheValPheLeuArgAsn-----PheSerLeuMetLeuArg 39
Db 251 AGGCTTTCCTCCGTCATCTTTTATTTAGGGGCCCACTTTAGTTGCTCCTT----- 198
QY 40 TyrTlGlyLysLysArgAlaThrArgPheTyrAspProArgArgGlyThrPro 58
Db 197 -----AGACCCCAAGAACTTTTCCCAAGAAAAATTGCCCC 156

RESULT 39
US-10-027-632-252580/c
Sequence 252580, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2005, 08:18:33 ; Search time 181 Seconds  
(without alignments)  
140.795 Million cell updates/sec

Title: US-10-031-158B-14

Perfect score: 58  
Sequence: 1 MGNFPPSPFLPFLQLKQSS.....RYIGKKRRARFMDPRRGTP 58

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 243163 seqs, 439378781 residues

Word size : 0

Total number of hits satisfying chosen parameters: 408321

Minimum DB seq length: 8  
Maximum DB seq length: 10

Post-processing: Listing first 45 summaries

Database :

1: Genesep21:\*  
2: genesep1980s:\*  
3: genesep2000s:\*  
4: genesep2001s:\*  
5: genesep2002s:\*  
6: genesep2003as:\*  
7: genesep2003bs:\*  
8: genesep2004s:\*  
9: genesep2005s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	15.5	9	ADM44402	Adm44402 Human TAR
2	9	15.5	9	ADM44401	Adm44401 Human TAR
3	9	15.5	9	ADM44406	Adm44406 Human TAR
4	8	13.8	8	ADM44405	Adm44405 Human TAR
5	8	13.8	9	ADM44404	Adm44404 Human TAR
6	6	10.3	9	ADH89809	Adh89809 Cell gene
7	6	10.3	9	ADM44403	Adm44403 Human TAR
8	5	8.6	8	AAH06282	AAH06282 p10SKI-1
9	5	8.6	9	AEA41508	AEA41508 GAG-blndi
10	5	8.6	9	AAH59128	AAH59128 Peptide f
11	5	8.6	9	AAH38039	AAH38039 Hepatitis
12	5	8.6	9	AAH38106	AAH38106 Hepatitis
13	5	8.6	9	AAH44612	AAH44612 Anti-fung
14	5	8.6	9	AAH43777	AAH43777 Bacterici
15	5	8.6	9	AAH45608	AAH45608 Immunogen
16	5	8.6	9	AAH45674	AAH45674 Immunogen
17	5	8.6	9	AAH46768	AAH46768 Immunogen
18	5	8.6	9	AAH00589	AAH00589 Anti-fung
19	5	8.6	9	AAH65513	AAH65513 Anti-fung
20	5	8.6	9	AAU00796	AAU00796 HIV-1 Tat
21	5	8.6	9	ABH08361	ABH08361 Synthetic
22	5	8.6	9	AAH015791	AAH015791 Human imm
23	5	8.6	9	ABP54737	ABP54737 HIV-1 tat
24	5	8.6	9	AAH35004	AAH35004 Mouse imm

25	5	8.6	9	ABH82213	ABH82213 Human ant
26	5	8.6	9	ADK65211	ADK65211 Human PIG
27	5	8.6	9	ADH97590	ADH97590 Immunogen
28	5	8.6	9	ADH89810	ADH89810 Cell gene
29	5	8.6	9	ADH44385	ADH44385 INK4a imm
30	5	8.6	9	ADH44376	ADH44376 INK4a imm
31	5	8.6	9	ADK84035	ADK84035 Human 191
32	5	8.6	9	ADK84519	ADK84519 Human 191
33	5	8.6	9	ADK85026	ADK85026 Human 191
34	5	8.6	9	ADK85074	ADK85074 Human 191
35	5	8.6	9	ADK87550	ADK87550 Human 191
36	5	8.6	9	ADK88447	ADK88447 Human 191
37	5	8.6	9	ADK83950	ADK83950 Human 191
38	5	8.6	9	ADK88445	ADK88445 Human 191
39	5	8.6	9	ADK84560	ADK84560 Human 191
40	5	8.6	9	ADK85620	ADK85620 Human 191
41	5	8.6	9	ADK84518	ADK84518 Human 191
42	5	8.6	9	ADK86086	ADK86086 Human 191
43	5	8.6	9	ADK86641	ADK86641 Human 191
44	5	8.6	9	ADK87058	ADK87058 Human 191
45	5	8.6	9	ADK87549	ADK87549 Human 191

#### ALIGNMENTS

RESULT 1  
ID ADM44402 standard; peptide; 9 AA.  
XX  
AC ADM44402;  
XX  
DT 24-MAR-2005 (first entry)  
XX  
DE Human TARP polypeptide epitope TARP-27-35.  
XX  
KW T-cell receptor gamma alternate reading frame protein; TARP;  
KW immunogenicity; breast tumor; prostate tumor; cytostatic; neoplasm;  
KW receptor.  
XX  
OS Homo sapiens.  
XX  
PN W02005000889-A1.  
XX  
PD 06-JAN-2005.  
XX  
PF 02-JUN-2004; 2004MO-US017574.  
XX  
PR 05-JUN-2003; 2003US-0476467P.  
XX  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX Berzofsky JA, Oh S, Pastan I;  
XX WPI; 2005-091494/10.  
XX  
XX New T-cell receptor gamma alternate reading frame protein (TARP)  
XX polypeptide, useful for treating a subject having breast or prostate  
XX cancer and for generating an immune response to TARP-expressing breast  
XX and prostate cancer cells.  
XX  
XX Claim 2; SEQ ID NO 4; 83pp; English.  
XX  
XX The invention relates to an immunogenic T-cell receptor gamma alternate  
XX reading frame protein (TARP) polypeptide and the polynucleotide encoding  
XX it. The invention also relates to a vector comprising the TARP  
XX polynucleotide, a host cell transformed with the vector, a pharmaceutical  
XX composition comprising a therapeutic amount of the polypeptide or the  
XX polynucleotide in a pharmaceutical carrier, eliciting an immune response  
XX in a subject, inhibiting the growth of a breast cancer or a prostate  
XX cancer cell, a reagent comprising a tetramer of the polypeptide bound to  
XX HLA-A2 and streptavidin, where the reagent is labeled or unlabeled, and  
XX detecting T cells expressing CD8 that specifically recognize the TARP

polypeptide in a subject. Elliciting an immune response in a subject comprises administering the polypeptide or polynucleotide, thus producing an immune response in the subject. The immune response comprises a T cell response or inducing cytotoxic T cells that induce lysis of cells expressing the TARP polypeptide. The subject has breast cancer or prostate cancer. The immune response decreases the growth of the prostate cancer or breast cancer. The method further comprises administering an adjuvant to the subject, inhibiting the growth of a breast cancer or prostate cancer cell comprises culturing cytotoxic T lymphocytes (CTLs) or CTL precursor cells with the polypeptide and an antigen presenting cell to produce activated CTLs matured from the CTL precursors that recognize the breast cancer or the prostate cancer cells, and contacting the breast cancer or the prostate cancer cell with the activated CTLs or CTLs matured from the CTL precursors, thus inhibiting the growth of the breast cancer or the prostate cancer cell. The TARP polypeptide and polynucleotide are useful for treating a subject having breast or prostate cancer. The polypeptide is useful for generating an immune response to breast cancer and prostate cancer cells that express TARP polypeptides. This sequence represents a human TARP polypeptide epitope used in the scope of the invention.

Query Match	15.5%	Score 9	DB 9	Length 9
Best Local Similarity	100.0%	Pred. No.	2e+06	
Matches 9	Conservative 0	Mismatches 0	Indels 0	Gaps 0

Qy	27	FVFLRNFSL	35
Db	1	FVFLRNFSL	9

```
RESULT 2
ADW44401
ID      ADW44401 standard; peptide; 9 AA.
```

DT 24-MAR-2005 (first entry)

DE Human TARP polypeptide epitope TARP-29-37.

KM T-cell receptor gamma alternate reading frame protein, TARP;  
 KM immunogenicity; breast tumor; prostate tumor; cytostatic; neoplasm,  
 KM receptor.

**Homo sapiens.**

PN WO2005000889-A1

PD 06-JAN-2005.

PF 02--JUN-2004; 2004WO-US017574.

PR 05-JUN-2003; 2003US-0476467P.

PA (USSH ) US DEPT HEALTH &amp; HUMAN SERVICES.

PI Berzofsky JA, Oh S, Pastan I;

DR WPI; 2005-091494/10.

PT New T-cell receptor gamma alternate reading frame protein (TARP)

PT cancer and for generating an immune response to TARP-expressing breast

XX

1. The first step is to identify the problem or question that needs to be addressed. This involves understanding the context and the specific requirements of the task.

CC reading frame protein (TARP) polypeptide and the polynucleotide encoding

CC polynucleotide, a host cell transformed with the vector, a pharmaceutical

CC composition comprising a therapeutic amount of the polypeptide or the  
CC polynucleotide in a pharmaceutical carrier, eliciting an immune response  
CC in a subject, inhibiting the growth of a breast cancer or a prostate  
CC cancer cell, a reagent comprising a tetramer of the polypeptide bound to  
CC HLA-A2 and streptavidin, where the reagent is labeled or unlabeled, and  
CC detecting T cells expressing CD8 that specifically recognize the TARP  
CC polypeptide in a subject. Eliciting an immune response in a subject  
CC comprises administering the polypeptide or polynucleotide, thus producing  
CC an immune response in the subject. The immune response comprises a T cell  
CC response or inducing cytotoxic T cells that induce lysis of cells  
CC expressing the TARP polypeptide. The subject has breast cancer or  
CC prostate cancer. The immune response decreases the growth of the prostate  
CC cancer or breast cancer. The method further comprises administering an  
CC adjuvant to the subject. Inhibiting the growth of a breast cancer or  
CC prostate cancer cell comprises culturing cytotoxic T lymphocytes (CTLs)  
CC or CTL precursor cells with the polypeptide and an antigen presenting  
CC cell to produce activated CTLs matured from the CTL precursors that  
CC recognize the breast cancer or the prostate cancer cells, and contacting  
CC the breast cancer or the prostate cancer cell with the activated CTLs or  
CC CTLs matured from the CTL precursors, thus inhibiting the growth of the  
CC breast cancer or the prostate cancer cell. The TARP polypeptide and  
CC polynucleotide are useful for treating a subject having breast or  
CC prostate cancer. The polypeptide is useful for generating an immune  
CC response to breast cancer and prostate cancer cells that express TARP  
CC polypeptides. This sequence represents a human TARP polypeptide epitope  
CC used in the scope of the invention

**SQ Sequence 9 AA**

Query Match	15.5%	Score 9;	DB 9;	Length 9;
Best Local Similarity	100.0%	Pred. No.	2e+06;	
Matches 9; Conservative	0;	Mismatches	0;	Indels 0;
				Gaps 0;

QY	29	PLRNFSML	3
Db	1	PLRNFSML	9

```

RESULT 3
ADW44406
ID ADW44406 standard; peptide; 9 AA

```

AC	ADW44406;
XX	
DT	24-MAR-2005 (first entry)

Human TARP polypeptide epitope TARP-22-30.

KW T-cell receptor gamma alternate reading frame protein; TARP;

KW receptor.

OS Homo sapiens.

PN WO2005000889-A1.

PD 06-JAN-2005.

PF 02-JUN-2004; 2004WO-US017574.

PR 05-JUN-2003; 2003US-0476467P.

PA (USSH ) US DEPT HEALTH &amp; HUMAN

PI Berzofsky JA, Oh S, Pastan I;

DR WPI; 2005-091494/10.

PT New T-cell receptor gamma alternate reading frame protein (TARP)

PT cancer and for generating an immune response to TARP-expressing breast

XX

PS Example 2; SEQ ID NO 8; 83bp; English.

XX The invention relates to an immunogenic T-cell receptor gamma alternate  
 CC reading frame protein (TARP) polypeptide and the polynucleotide encoding  
 CC it. The invention also relates to a vector comprising the TARP  
 CC polynucleotide, a host cell transformed with the vector, a pharmaceutical  
 CC composition comprising a therapeutic amount of the polypeptide or the  
 CC polynucleotide in a pharmaceutical carrier, eliciting an immune response  
 CC in a subject, inhibiting the growth of a breast cancer or a prostate  
 CC cancer cell, a reagent comprising a tetramer of the polypeptide bound to  
 CC HLA-A2 and streptavidin, where the reagent is labeled or unlabeled, and  
 CC detecting T cells expressing CD8 that specifically recognize the TARP  
 CC polypeptide in a subject. Eliciting an immune response in a subject  
 CC comprises administering the polypeptide or polynucleotide, thus producing  
 CC an immune response in the subject. The immune response comprises a T cell  
 CC response or inducing cytotoxic T cells that induce lysis of cells  
 CC expressing the TARP polypeptide. The subject has breast cancer or  
 CC prostate cancer. The immune response decreases the growth of the prostate  
 CC cancer or breast cancer. The method further comprises administering an  
 CC adjuvant to the subject. Inhibiting the growth of a breast cancer or  
 CC prostate cancer cell comprises culturing cytotoxic T lymphocytes (CTLs)  
 CC or CTL precursor cells with the polypeptide and an antigen presenting  
 CC cell to produce activated CTLs matured from the CTL precursors that  
 CC recognize the breast cancer or the prostate cancer cells, and contacting  
 CC the breast cancer or the prostate cancer cell with the activated CTLs or  
 CC CTLs matured from the CTL precursors, thus inhibiting the growth of the  
 CC breast cancer or the prostate cancer cell. The TARP polypeptide and  
 CC polynucleotide are useful for treating a subject having breast or  
 CC prostate cancer. The polypeptide is useful for generating an immune  
 CC response to breast cancer and prostate cancer cells that express TARP  
 CC polypeptides. This sequence represents a human TARP polypeptide epitope  
 CC used in the scope of the invention.

XX Sequence 9 AA;

SQ

Query Match 15.5%; Score 9; DB 9; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 2e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 RLHETPVFL 30  
 |||||  
 1 RLHETPVFL 9

Db

RESULT 4  
 ADM44405  
 ID ADM44405 standard; peptide; 8 AA.  
 XX  
 AC ADM44405;  
 XX  
 XX 24-MAR-2005 (first entry)  
 XX  
 DE Human TARP polypeptide epitope TARP-2-9.  
 XX  
 XX T-cell receptor gamma alternate reading frame protein; TARP;  
 KM immunogenicity; breast tumor; prostate tumor; cytostatic; neoplasm;  
 KM receptor.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX WO2005000889-A1.  
 PN  
 PD 06-JAN-2005.  
 XX  
 PF 02-JUN-2004; 2004WO-US017574.  
 XX  
 PR 05-JUN-2003; 2003US-0476467P.  
 XX  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA  
 XX Bezofeky JA, Oh S, Pastan I;  
 PI  
 XX WPI, 2005-091494/10.

XX New T-cell receptor gamma alternate reading frame protein (TARP)  
 PT polypeptide, useful for treating a subject having breast or prostate  
 PT cancer and for generating an immune response to TARP-expressing breast  
 PT and prostate cancer cells.

XX Example 2; SEQ ID NO 7; 83bp; English.

XX The invention relates to an immunogenic T-cell receptor gamma alternate  
 CC reading frame protein (TARP) polypeptide and the polynucleotide encoding  
 CC it. The invention also relates to a vector comprising the TARP  
 CC polynucleotide, a host cell transformed with the vector, a pharmaceutical  
 CC composition comprising a therapeutic amount of the polypeptide or the  
 CC polynucleotide in a pharmaceutical carrier, eliciting an immune response  
 CC in a subject, inhibiting the growth of a breast cancer or a prostate  
 CC cancer cell, a reagent comprising a tetramer of the polypeptide bound to  
 CC HLA-A2 and streptavidin, where the reagent is labeled or unlabeled, and  
 CC detecting T cells expressing CD8 that specifically recognize the TARP  
 CC polypeptide in a subject. Eliciting an immune response in a subject  
 CC comprises administering the polypeptide or polynucleotide, thus producing  
 CC an immune response in the subject. The immune response comprises a T cell  
 CC response or inducing cytotoxic T cells that induce lysis of cells  
 CC expressing the TARP polypeptide. The subject has breast cancer or  
 CC prostate cancer. The immune response decreases the growth of the prostate  
 CC cancer or breast cancer. The method further comprises administering an  
 CC adjuvant to the subject. Inhibiting the growth of a breast cancer or  
 CC prostate cancer cell comprises culturing cytotoxic T lymphocytes (CTLs)  
 CC or CTL precursor cells with the polypeptide and an antigen presenting  
 CC cell to produce activated CTLs matured from the CTL precursors that  
 CC recognize the breast cancer or the prostate cancer cells, and contacting  
 CC the breast cancer or the prostate cancer cell with the activated CTLs or  
 CC CTLs matured from the CTL precursors, thus inhibiting the growth of the  
 CC breast cancer or the prostate cancer cell. The TARP polypeptide and  
 CC polynucleotide are useful for treating a subject having breast or  
 CC prostate cancer. The polypeptide is useful for generating an immune  
 CC response to breast cancer and prostate cancer cells that express TARP  
 CC polypeptides. This sequence represents a human TARP polypeptide epitope  
 CC used in the scope of the invention.

XX Sequence 8 AA;

SQ

Query Match 13.8%; Score 8; DB 9; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 2e+06;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QMFPPSPFL 9  
 |||||  
 1 QMFPPSPFL 8

Db

RESULT 5  
 ADM44404  
 ID ADM44404 standard; peptide; 9 AA.  
 XX  
 AC ADM44404;  
 XX  
 XX 24-MAR-2005 (first entry)  
 XX  
 DE Human TARP polypeptide epitope TARP-29-37-9V.  
 XX  
 XX T-cell receptor gamma alternate reading frame protein; TARP;  
 KM immunogenicity; breast tumor; prostate tumor; cytostatic; neoplasm;  
 KM receptor; mutain.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX Synthetic.  
 OS  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 9 /note= "Wild-type Leu substituted by Val"  
 FT  
 XX  
 PN WO2005000889-A1.

PD 06-JAN-2005.  
 XX 02-JUN-2004; 2004WO-US017574.  
 PF  
 XX 05-JUN-2003; 2003US-0476467P.  
 PR  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA Berzofsky JA, Oh S, Pastan I;  
 PI  
 XX WPI: 2005-091494/10.  
 DR  
 XX  
 PT New T-cell receptor gamma alternate reading frame protein (TARP)  
 PT polypeptide, useful for treating a subject having breast or prostate  
 PT cancer and for generating an immune response to TARP-expressing breast  
 PT and prostate cancer cells.  
 PT  
 XX  
 PS Claim 2; SEQ ID NO 6; 83pp; English.  
 XX  
 CC The invention relates to an immunogenic T-cell receptor gamma alternate  
 CC reading frame protein (TARP) polypeptide and the polynucleotide encoding  
 CC it. The invention also relates to a vector comprising the TARP  
 CC polynucleotide, a host cell transformed with the vector, a pharmaceutical  
 CC composition comprising a therapeutic amount of the polypeptide or the  
 CC polynucleotide in a pharmaceutical carrier, eliciting an immune response  
 CC in a subject, inhibiting the growth of a breast cancer or a prostate  
 CC cancer cell, a reagent comprising a tetramer of the polypeptide bound to  
 CC HLA-A2 and streptavidin, where the reagent is labeled or unlabeled, and  
 CC detecting T cells expressing CD8 that specifically recognize the TARP  
 CC polypeptide in a subject. Eliciting an immune response in a subject  
 CC comprises administering the polypeptide or polynucleotide, thus producing  
 CC an immune response in the subject. The immune response comprises a T cell  
 CC response or inducing cytotoxic T cells that induce lysis of cells  
 CC expressing the TARP polypeptide. The subject has breast cancer or  
 CC prostate cancer. The immune response decreases the growth of the prostate  
 CC cancer or breast cancer. The method further comprises administering an  
 CC adjuvant to the subject. Inhibiting the growth of a breast cancer or  
 CC prostate cancer cell comprises culturing cytotoxic T lymphocytes (CTLs)  
 CC or CTL precursor cells with the polypeptide and an antigen presenting  
 CC cell to produce activated CTLs matured from the CTL precursors that  
 CC recognize the breast cancer or the prostate cancer cells, and contacting  
 CC the breast cancer or the prostate cancer cell with the activated CTLs or  
 CC CTLs matured from the CTL precursors, thus inhibiting the growth of the  
 CC breast cancer or the prostate cancer cell. The TARP polypeptide and  
 CC polynucleotide are useful for treating a subject having breast or  
 CC prostate cancer. The polypeptide is useful for generating an immune  
 CC response to breast cancer and prostate cancer cells that express TARP  
 CC polypeptides. This sequence represents a human TARP polypeptide epitope  
 CC used in the scope of the invention.  
 CC  
 XX Sequence 9 AA:  
 SQ  
 Query Match 13.8%; Score 8; DB 9; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 2e+06;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 29 FLRNFSLM 36  
 Db 1 FLRNFSLM 8  
 RESULT 6  
 ADH89809  
 ID ADH89809 standard; peptide; 9 AA.  
 AC  
 XX ADH89809;  
 DT  
 XX 22-APR-2004 (first entry)  
 DE Cell penetrating peptide (CPP) identification method-related peptide 11.  
 XX  
 XX cell-penetrating peptide; CPP; bulk property value Z-E; Z-E1; Z-E2; Z-E3;  
 KW Z-E4; Z-E5; antidiabetic; neuroprotective; nootropic; antiparkinsonian;

KW cardiant; cytostatic; tranquiliser; immunosuppressive; antidepressant;  
 KW anticonvulsant; antiinflammatory; analgesic; neuroleptic;  
 KW ophthalmological; antitumor; cell-penetrating; infectious disease;  
 KW diabetes type II; diabetes type II; Alzheimer's disease;  
 KW Parkinson's disease; cancer; prion disease; cardiovascular disease;  
 KW signal transduction.  
 XX  
 XX Unidentified.  
 OS  
 XX WO2003106491-A2.  
 XX  
 XX 24-DEC-2003.  
 XX  
 PD  
 XX 18-JUN-2003; 2003WO-IB003163.  
 PF 18-JUN-2002; 2002SE-00001863.  
 XX 25-JUN-2002; 2002US-0391788P.  
 PR  
 XX (CEPE-) CEPEP AB.  
 PA  
 XX Haeilbrink M, Pooga M, Metsis M, Kogerman P, Valkna A, Melkas A;  
 PI Lindgren M, Graesslund A, Eriksson G, Oestensson CG, Budhina M;  
 PI Zorko M, Elmquist A, Soomets U, Lundberg P, Jaerwer P, Saar K;  
 PI El-Andalousi S, Kilik K, Langel U;  
 XX  
 DR WPI: 2004-090832/09.  
 XX  
 XX  
 PT Predicting, designing, detecting, and/or verifying novel cell-penetrating  
 PT peptide based on assessment of bulk property value of sequences of cell-  
 PT penetrating peptide.  
 PT  
 XX Disclosure; Page 31; 148pp; English.  
 XX  
 PS  
 CC This invention relates to a novel method of identifying, designing,  
 CC detecting, and/or verifying novel cell-penetrating peptide (CPP) based on  
 CC assessment of bulk property value Z-E of sequences of CPP comprising 5 or  
 CC more individual average interval values Z-E1, Z-E2, Z-E3, Z-E4 and Z-E5,  
 CC where Z-E1, Z-E2, Z-E3, Z-E4 and Z-E5 are average values of the  
 CC respective descriptor values for the residues in the amino acid sequence.  
 CC The invention may be useful for the development of compounds with an  
 CC antidiabetic, neuroprotective, nootropic, antiparkinsonian, cardiant,  
 CC cyostatic, tranquiliser, immunosuppressive, antidepressant,  
 CC anticonvulsant, antiinflammatory, analgesic, neuroleptic,  
 CC ophthalmological or antitumor activity as a stimulator of cell-  
 CC penetration. The method of the invention is useful for identifying a cell  
 CC -penetrating peptide or protein and/or a cell-penetrating fragment of a  
 CC peptide or protein. In addition, the invention may be useful for checking  
 CC cellular penetration properties of a peptide, for producing a cell-  
 CC penetrating and functional protein-mimicking peptide and for de novo  
 CC design and production of an artificial cell-penetrating and/or and  
 CC artificial cell-penetrating and functional protein-mimicking peptide.  
 CC Compositions developed within the scope of the present invention may be  
 CC useful for treating infectious diseases, diabetes type I, diabetes type  
 CC II, Alzheimer's disease, Parkinson's disease, cancer, prion disease,  
 CC cardiovascular disease or disorders resulting from perturbed signal  
 CC transduction. The method of the invention is fast, efficient and reliable  
 CC for identifying, detecting, designing CPPs and for screening cellular  
 CC uptake of a broad variety of CPPs in vitro and in vivo. The present  
 CC sequence is that of a peptide which is related to the invention.  
 CC  
 XX Sequence 9 AA:  
 SQ  
 Query Match 10.3%; Score 6; DB 8; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 2e+06;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 12 FLQLLK 17  
 Db 4 FLQLLK 9  
 RESULT 7  
 ADW44403

ID ADM44403 standard; peptide: 9 AA.  
 AC ADM44403;  
 XX  
 XX  
 DT 24-MAR-2005 (first entry)  
 DE Human TARP polypeptide epitope TARP-29-37-3A.  
 XX  
 XX T-cell receptor gamma alternate reading frame protein; TARP;  
 KM immunogenicity; breast tumor; prostate tumor; cytostatic; neoplasm;  
 KM receptor; mutuin.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 3 /note="Wild-type Arg substituted by Ala"  
 FT  
 FN WO2005000889-A1.  
 XX  
 PD 06-JAN-2005.  
 XX  
 PF 02-JUN-2004; 2004WO-US017574.  
 XX  
 PR 05-JUN-2003; 2003US-0476467P.  
 XX  
 PA (US\$ ) US DEPT HEALTH & HUMAN SERVICES.  
 PI Berzofsky JA, Oh S, Pastan I;  
 DR WPI; 2005-091494/10.  
 XX  
 PT New T-cell receptor gamma alternate reading frame protein (TARP)  
 PT polypeptide useful for treating a subject having breast or prostate  
 PT cancer and for generating an immune response to TARP-expressing breast  
 PT and prostate cancer cells.  
 PT  
 PS Claim 2; SEQ ID NO 5; 83pp; English.  
 XX  
 CC The invention relates to an immunogenic T-cell receptor gamma alternate  
 CC reading frame protein (TARP) polypeptide and the polynucleotide encoding  
 CC it. The invention also relates to a vector comprising the TARP  
 CC polynucleotide, a host cell transformed with the vector, a pharmaceutical  
 CC composition comprising a therapeutic amount of the polypeptide or the  
 CC polynucleotide in a pharmaceutical carrier, eliciting an immune response  
 CC in a subject, inhibiting the growth of a breast cancer or a prostate  
 CC cancer cell, a reagent comprising a tetramer of the polypeptide bound to  
 CC HLA-A2 and streptavidin, where the reagent is labeled or unlabeled, and  
 CC detecting T cells expressing CD8 that specifically recognize the TARP  
 CC polypeptide in a subject. Eliciting an immune response in a subject  
 CC comprises administering the polypeptide or polynucleotide, thus producing  
 CC an immune response in the subject. The immune response comprises a T cell  
 CC response or inducing cytotoxic T cells that induce lysis of cells  
 CC expressing the TARP polypeptide. The subject has breast cancer or  
 CC prostate cancer. The immune response decreases the growth of the prostate  
 CC cancer or breast cancer. The method further comprises administering an  
 CC adjuvant to the subject. Inhibiting the growth of a breast cancer or  
 CC prostate cancer cell comprises culturing cytotoxic T lymphocytes (CTLs)  
 CC or CTL precursor cells with the polypeptide and an antigen presenting  
 CC cell to produce activated CTLs matured from the CTL precursors that  
 CC recognize the breast cancer or the prostate cancer cells, and contacting  
 CC the breast cancer or the prostate cancer cell with the activated CTLs or  
 CC CTLs matured from the CTL precursors, thus inhibiting the growth of the  
 CC breast cancer or the prostate cancer cell. The TARP polypeptide and  
 CC polynucleotide are useful for treating a subject having breast or  
 CC prostate cancer. The polypeptide is useful for generating an immune  
 CC response to breast cancer and prostate cancer cells that express TARP  
 CC polypeptides. This sequence represents a human TARP polypeptide epitope  
 CC used in the scope of the invention.  
 CC  
 XX Sequence 9 AA;  
 SQ

Query Match 10.3%; Score 6; DB 9; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 2e+06;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 32 NFSJML 37  
 DB 4 NFSJML 9

RESULT 8  
 AAB06282  
 ID AAB06282 standard; peptide: 8 AA.  
 AC AAB06282;  
 XX  
 DT 03-OCT-2000 (first entry)  
 DE proSKI-1 putative zymogen cleavage site.  
 XX  
 XX Human; pro-brain-derived neurotrophic factor; proBDNF;  
 KM subtilisin-kexin isoenzyme 1; SKI-1; proSKI-1 zymogen processing;  
 KM antihypertensive; cytostatic; vasotropic; SKI-1 inhibitor;  
 KM hypercholesterolemia; liver steatosis; Ras-dependent cancer; restenosis;  
 KM amyloid protein formation.  
 XX  
 OS Homo sapiens.  
 OS Mus sp.  
 OS Rattus sp.  
 XX  
 FH Key Location/Qualifiers  
 FT Cleavage-site 4..5  
 FT  
 FN WO200026348-A2.  
 XX  
 PD 11-MAY-2000.  
 XX  
 PF 04-NOV-1999; 99WO-CA001058.  
 XX  
 PR 04-NOV-1998; 98CA-02249648.  
 XX

PA (RECL-) INST RECH CLINIQUES MONTREAL.  
 PI Seidah N, Chretien M, Marcinkiewicz M, Laaksonen R, Davignon J;  
 DR WPI; 2000-365601/31.  
 XX

PT Novel soluble proteic fragment of subtilisin-kexin isoenzyme for  
 PT producing a polypeptide useful for treating hypercholesterolemia, liver  
 PT steatosis and amyloidosis, comprises a specific amino acid sequence.  
 XX

PS Example 1; Page 25; 119pp; English.  
 XX

CC The present sequence is a possible cleavage site in proSKI-1 which may be  
 CC targeted by a zymogen to form mature subtilisin-kexin isoenzyme 1 (SKI-  
 CC 1). The conserved site was identified by alignment of SKI-1 with other  
 CC subtilases. SKI-1 is a type-1 membrane-bound proteinase which cleaves pro  
 CC -brain-derived neurotrophic factor. Peptides which bind to and are  
 CC cleaved by SKI-1 may be used for monitoring SKI-1 activity, for screening  
 CC inhibitors of SKI-1 activity, or for screening enhancers of SKI-1  
 CC activity. Proteic fragments of SKI-1 which bind to the SKI-1 catalytic  
 CC site may be used as inhibitors of SKI-1 activity. They may be used to  
 CC treat diseases involving overexpression of SKI-1 or SKI-1 substrate. Such  
 CC diseases include hypercholesterolemia, high levels of fatty acids, liver  
 CC lipid or farnesyl pyrophosphate, liver steatosis, Ras-dependent cancer,  
 CC restenosis and amyloid protein formation  
 CC  
 XX Sequence 8 AA;  
 SQ

Query Match 8.6%; Score 5; DB 3; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 2e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 19 SSRLL 23

Db 3 SSRRL 7

RESULT 9  
AEA41508  
ID AEA41508 standard; peptide; 8 AA.  
XX  
AC AEA41508;  
XX  
DT 25-AUG-2005 (first entry)  
XX  
DE GAG-binding peptide from FAS binding protein #2.  
XX  
KM Glycosaminoglycan; heparan sulfate; protein engineering; chemokine;  
KM inflammation; rheumatoid arthritis; psoriasis; osteoarthritis; asthma;  
KM Alzheimer's disease; multiple sclerosis; Antiinflammatory; Antiarthritic;  
KM Antirheumatic; Antipsoriatic; Osteopathic; Antiasthmatic;  
KM Neutroprotective; Nootropic.  
XX  
OS Unidentified.  
XX  
PN WO2005054265-A1.  
PD 16-JUN-2005.  
PF 02-DEC-2004; 2004WO-EP013670.  
XX  
PR 04-DEC-2003; 2003AT-00001952.  
XX  
PA (KUNGL) KUNGL A J.  
XX  
PI Kungl AJ;  
XX  
DR WPI; 2005-435363/44.  
PT Introducing a glycosaminoglycan (GAG) binding site into a protein, useful  
PT for producing mutant chemokines which disrupt chemokine signaling, and  
PT can be used to treat inflammatory diseases.  
XX  
XX Example 6; Page 38; 69pp; English.  
XX  
CC The invention relates to introducing a glycosaminoglycan (GAG) binding  
CC site into a protein comprising identifying a region in the protein which  
CC is not essential for structure maintenance, and introducing at least one  
CC basic amino acid and/or deleting at least one bulky and/or acidic amino  
CC acid in the identified region. The GAG binding site has a GAG binding  
CC affinity of Kd less than or equal to 10 microm, preferably less than or  
CC equal to 1 microm, still more preferably less than or equal to 0.1  
CC microm. Also included are a protein obtainable by the method, a modified  
CC GAG binding protein (A) (characterized in that a GAG binding region in  
CC the protein is modified by substitution, insertion, and/or deletion of at  
CC least one amino acid in order to increase the relative amount of basic  
CC amino acids in the GAG binding region, and/or reduce the amount of bulky  
CC and/or acidic amino acids in the GAG binding region, preferably at a  
CC solvent exposed position, and in that the GAG binding affinity of the  
CC protein is increased compared to the GAG binding affinity of a respective  
CC wild-type protein), an isolated polynucleic acid molecule encoding the  
CC modified protein, a vector comprising an isolated polynucleic acid  
CC molecule, a recombinant cell stably transfected with the vector and a  
CC pharmaceutical composition comprising a carrier and the protein,  
CC polynucleic acid molecule or vector. The protein is a chemokine,  
CC preferably IL-8, RANTES or MCP-1. The protein has increased binding  
CC affinity to heparan sulfate and/or heparin. A further biologically active  
CC region (e.g. for leukocyte activation) is modified, inhibiting or down-  
CC regulating the further biological activity of the protein. The protein,  
CC polynucleic acid, or vector is useful for inhibiting or suppressing the  
CC biological activity of the respective wild-type protein. The protein,  
CC polynucleic acid, or vector is useful for producing a medicament for  
CC treating an inflammatory condition, preferably a condition selected from  
CC rheumatoid arthritis, psoriasis, osteoarthritis, asthma, Alzheimer's  
CC disease, and multiple sclerosis. The method provides a means of blocking  
CC the interaction between wild-type chemokines and their glycosaminoglycan

CC (GAG) co-receptors, and represents a new strategy for blocking chemokine  
CC signaling. The present sequence is a GAG-binding site.  
XX  
SQ Sequence 8 AA;  
XX  
Query Match 8.6%; Score 5; DB 9; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 43 KKRRRA 47  
Db 2 KKRRRA 6  
XX  
RESULT 10  
AAR59128  
ID AAR59128 standard; peptide; 9 AA.  
XX  
AC AAR59128;  
XX  
DT 16-OCT-2003 (revised)  
DT 25-MAR-2003 (revised)  
DT 02-MAY-1995 (first entry)  
XX  
DE Peptide fragment (1.0882) of HBV binds HLA-A2.1.  
XX  
KM antigen; epitope; immunogenic target protein; PSA; HBVC; HBVs; EBV; HIV1;  
KM core antigen; surface antigen; pharmaceutical composition; in vivo;  
KM ex vivo; therapeutic; diagnostic; MHC class I molecule;  
KM major histocompatibility complex; HLA-A2.1; 9mer; 10mer; anchor;  
KM human leukocyte antigen.  
XX  
OS Hepatitis B virus; strain adw.  
XX  
PN WO9420127-A1.  
PD 15-SEP-1994.  
PF 04-MAR-1994; 94WO-US002353.  
XX  
PR 05-MAR-1993; 93US-00027146.  
PR 04-JUN-1993; 93US-00073205.  
PR 29-NOV-1993; 93US-00159184.  
XX  
PA (CYTE-) CYTEL CORP.  
XX  
PI Grey HM, Sette A, Sidney J, Kast W;  
XX  
DR WPI; 1994-302678/37.  
PT Immunogenic peptide(s) having an HLA-A2.1 binding motif - used for  
PT treatment or prophylaxis of cancer, virus infection or autoimmune  
PT diseases.  
XX  
PS Example 5; Page 101; 138pp; English.  
XX  
CC AAR59104-264 are immunogenic 9mer peptides that contain a HLA-A2.1  
CC binding motif. These peptides bind HLA-A2.1 and have a binding affinity  
CC of at least 1% as compared to a reference peptide (AAR71293). AAR59128  
CC has an IC50 of 0.046 and the sequence occurs at position 1088 in the HBV  
CC POL protein. The peptides of the invention can induce cytotoxic T  
CC lymphocytes which can react with target cells. They can be used for the  
CC treatment or prophylaxis of cancer, eg. prostate cancer or lymphoma, etc.  
CC (Updated on 25-MAR-2003 to correct PW field.) (Updated on 16-OCT-2003 to  
CC standardise OS field)  
XX  
SQ Sequence 9 AA;  
QY  
Query Match 8.6%; Score 5; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 2e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 34 SLMDLL 38

Db |||||  
5 SLMLL 9

## RESULT 11

AAV38039 ID AAV38039 standard; peptide; 9 AA.

AC AAV38039;

DT 29-SEP-1999 (first entry)

DE Hepatitis B virus-derived HLA-binding peptide.

XX Immunogen; HLA; human leukocyte antigen; binding motif; antiviral; MHC;  
KM major histocompatibility complex; viral infection; anticancer;  
KW prostate cancer; lymphoma; hepatitis; AIDS; diagnostic; diagnosis.

XX Hepatitis B virus.

OS WO9403205-A1.

PN 17-FEB-1994.

PD 06-AUG-1993; 93WO-US007421.

PF 07-AUG-1992; 92US-00926666.

PR 05-MAR-1993; 93US-00027746.

PS (CYTE-) CYTEL CORP.

PI Kubo RT, Grey HM, Sette A, Celis E;

DR WPI; 1994-065403/08.

XX Peptide which specifically binds selected MHC allele - used to induce an  
PT immune response for treatment or prevention of viral infection or cancer,  
PT or for diagnosis.

PS Disclosure; Page 106; 150pp; English.

XX The sequence is a specific example of a group of new immunogenic peptides  
CC having an HLA-A3.2, HLA-A1, HLA-A11 or HLA-A24.1 binding motif. For  
CC example, the peptides having an HLA-A3.2 binding motif each have 9-10  
CC residues and contain, from the N-terminus to the C-terminus, (a) a first  
CC conserved residue selected from L, M, I, V, S, A, T, F, C, G, D and E and  
CC (b) a second conserved residue from K, R, Y, H or F, where the first and  
CC second conserved residues are separated by 6-7 residues. The peptides are  
CC capable of binding selected MHC molecules and inducing an immune  
CC response. They can be used to treat and/or prevent viral infection and  
CC cancer, e.g. prostate cancer, lymphoma, hepatitis or AIDS. They can also  
CC be used to produce antibodies for use as diagnostic or therapeutic  
CC agents. The peptides can also be used as diagnostic agents

XX Sequence 9 AA:

Query Match 8.6%; Score 5; DB 2; Length 9;

Best Local Similarity 100.0%; Pred. No. 2e+06;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 34 SLMLL 38

DB 1 SLMLL 5

## RESULT 12

AAV38106 ID AAV38106 standard; peptide; 9 AA.

AC AAV38106;

DT 29-SEP-1999 (first entry)

DE Hepatitis B virus-derived HLA-binding peptide.

XX Immunogen; HLA; human leukocyte antigen; binding motif; antiviral; MHC;  
KM major histocompatibility complex; viral infection; anticancer;  
KW prostate cancer; lymphoma; hepatitis; AIDS; diagnostic; diagnosis.

XX Hepatitis B virus.

OS WO9403205-A1.

PN 17-FEB-1994.

PD 06-AUG-1993; 93WO-US007421.

PF 07-AUG-1992; 92US-00926666.

PR 05-MAR-1993; 93US-00027746.

PS (CYTE-) CYTEL CORP.

PI Kubo RT, Grey HM, Sette A, Celis E;

DR WPI; 1994-065403/08.

XX Peptide which specifically binds selected MHC allele - used to induce an  
PT immune response for treatment or prevention of viral infection or cancer,  
PT or for diagnosis.

PS Disclosure; Page 107; 150pp; English.

XX The sequence is a specific example of a group of new immunogenic peptides  
CC having an HLA-A3.2, HLA-A1, HLA-A11 or HLA-A24.1 binding motif. For  
CC example, the peptides having an HLA-A3.2 binding motif each have 9-10  
CC residues and contain, from the N-terminus to the C-terminus, (a) a first  
CC conserved residue selected from L, M, I, V, S, A, T, F, C, G, D and E and  
CC (b) a second conserved residue from K, R, Y, H or F, where the first and  
CC second conserved residues are separated by 6-7 residues. The peptides are  
CC capable of binding selected MHC molecules and inducing an immune  
CC response. They can be used to treat and/or prevent viral infection and  
CC cancer, e.g. prostate cancer, lymphoma, hepatitis or AIDS. They can also  
CC be used to produce antibodies for use as diagnostic or therapeutic  
CC agents. The peptides can also be used as diagnostic agents

XX Sequence 9 AA:

Query Match 8.6%; Score 5; DB 2; Length 9;

Best Local Similarity 100.0%; Pred. No. 2e+06;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 34 SLMLL 38

DB 3 SLMLL 7

## RESULT 13

AAW44612 ID AAW44612 standard; peptide; 9 AA.

AC AAW44612;

DT 27-APR-1998 (first entry)

DE Anti-fungal peptide #213 based on BPI protein (residues 142-169).

XX Anti-fungal peptide; bactericidal-permeability-increasing protein; BPI;

KW polymorphonuclear leukocyte; fungicide.

XX Synthetic.

OS Mammalia.

XX Key

FT Modified-site 9 Location/Qualifiers

FT /note= "C-terminal amide"

PN W09704008-A1.  
 XX  
 PD 06-FEB-1997.  
 XX  
 PF 21-MAR-1996; 96WO-US003845.  
 XX  
 PR 20-JUL-1995; 95US-00504841.  
 XX  
 PA (XOMA ) XOMA CORP.  
 XX  
 PI Little RG, Lim E, Fadem MB;  
 XX  
 DR WPI; 1997-132578/12.  
 XX  
 XX Anti-fungal peptide(s) derived from or based on domain III of  
 PT bactericidal/permeability-increasing protein - are used in vitro or in  
 CC vivo as a fungicides.  
 XX  
 PS Claim 1; -pp; 230pp; English.  
 XX  
 CC This is a specifically claimed anti-fungal peptide which is based on  
 CC domain III (amino acids 142-160) of bactericidal-permeability-increasing  
 CC protein (BPI), isolated from the granules of mammalian polymorphonuclear  
 CC leukocytes. It is used in compositions with diluents, carriers or  
 CC adjuvants to treat fungal infections in patients. It may also be used in  
 CC vitro to kill or inhibit the replication of fungi, such as in  
 CC decontaminating fluids and sterilising medical and implant devices  
 CC  
 SQ Sequence 9 AA;

Query Match 8.6%; Score 5; DB 2; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 2e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LQLLK 17  
 |||||  
 Db 4 LQLLK 8

RESULT 14  
 AAM43777  
 ID AAM43777 standard; peptide; 9 AA.  
 XX

AC AAM43777;

XX 20-APR-1998 (first entry)

DE Bactericidal/permeability increasing peptide XMP.382.

XX Bactericidal/permeability increasing peptide; BPI; fusion protein;

KW bacterial infection; fungal infection; endotoxin; heparin; angiogenesis;

XX fungicidal; recombinant DNA; vector.

OS Homo sapiens.

OS Synthetic.

XX Key

XX Modified-site

XX Location/Qualifiers

XX 9 /note= "Amdated"

XX W09735009-A1.

XX 25-SEP-1997.

XX 18-MAR-1997; 97WO-US005287.

XX 22-MAR-1996; 96US-00621803.

XX (XOMA ) XOMA CORP.

XX Better MD;

XX WPI; 1997-480215/44.

XX  
 PT Recombinant production of bactericidal/permeability increasing protein -  
 PT by expression as a fusion protein in microbial host cells, then cleaving  
 PT the BPI peptide from the carrier.  
 XX  
 PS Claim 10; Page 136; 186pp; English.

XX  
 CC A new recombinant DNA vector construct has been developed which encodes a  
 CC fusion protein and is suitable for introduction into a bacterial host.  
 CC The vector comprises: (a) DNA encoding at least one cationic  
 CC bactericidal/permeability increasing peptide (BPI), (b) DNA encoding a  
 CC carrier protein, and (c) DNA encoding an amino acid (aa) cleavage site  
 CC located between (a) and (b). The present sequence represents a  
 CC specifically claimed BPI peptide. The peptides have many uses including  
 CC the treatment of bacterial and fungal infections. BPI peptides also bind  
 CC to endotoxins and heparin, neutralising their effects. The peptides have  
 CC further been shown to inhibit angiogenesis (partly due to heparin-binding  
 CC activity). The fusion proteins have been found to be expressed in large  
 CC amounts without significant proteolysis, and in some cases are actually  
 CC secreted from the host cells. This allows the indirect production of anti-  
 CC microbial BPI peptides in microbial hosts

SQ Sequence 9 AA;

Query Match 8.6%; Score 5; DB 2; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 2e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LQLLK 17  
 |||||  
 Db 4 LQLLK 8

RESULT 15  
 AAY45608  
 ID AAY45608 standard; peptide; 9 AA.  
 XX

AC AAY45608;

XX 01-DEC-1999 (first entry)

DE Immunogenic peptide having a human leukocyte antigen binding motif #219.

XX Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;

KW immune response; T cell activation; major histocompatibility complex;

KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;

KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;

KW vaccine; immunisation.

XX Synthetic.

OS Homo sapiens.

XX W09945954-A1.

XX 16-SEP-1999.

XX 13-MAR-1998; 98WO-US005039.

XX 13-MAR-1998; 98WO-US005039.

XX (EPIIM-) EPIMMUNE INC.

XX Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;

XX WPI; 1999-551214/46.

XX New immunogenic peptides with HLA binding motif, useful in treatment and

XX diagnosis of cancers and viral diseases.

XX Claim 1; Page 37; 150pp; English.

XX AAY45390 to AAY48214 represent specifically claimed immunogenic peptides

CC having a human major histocompatibility complex (MHC) Class I (also known

CC as human leukocyte antigen (HLA)) binding motif. The immunogenic peptides  
 CC can bind to a specific HLA allele (i.e. HLA-A subtypes HLA-A2.1, A1, A3.2  
 CC or A24.1 or HLA-B or C) and induce a cytotoxic T cell response against  
 CC the antigen from which the peptide is derived. Cytotoxic T lymphocytes  
 CC (CTLs) which destroy antigen-bearing cells are normally induced by an  
 CC antigen in the form of a peptide fragment bound to a HLA molecule, rather  
 CC than the intact foreign antigen itself, and are particularly important in  
 CC tumour rejection and in fighting viral infections. The peptides are  
 CC therefore useful therapeutically to treat or prevent viral infections and  
 CC cancers in mammals (especially humans) e.g. prostate cancer, hepatitis B  
 CC and C, AIDS, and renal carcinoma. They can be administered as vaccines to  
 CC elicit an immune response in individuals susceptible or otherwise at risk  
 CC of viral infection or cancer, or used to treat chronic or acute  
 CC conditions. They are also useful diagnostically, and can be used to  
 CC induce a cytotoxic T cell response, by contacting a cytotoxic T cell with  
 CC the peptide e.g. to produce CTLs ex vivo for infusion back into a  
 CC patient. The polynucleotides encoding the immunogenic peptides are also  
 CC useful therapeutically and for immunisation as above  
 CC  
 XX

SQ Sequence 9 AA;

Query Match

Best Local Similarity 8.6%; Score 5; DB 2; Length 9;

Best Local Similarity 100.0%; Pred. No. 2e+06;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 SLML 38

Db 1 SLML 5

Search completed: December 13, 2005, 08:21:44  
 Job time : 183 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2005, 08:18:36 ; Search time 37 Seconds

(without alignments)  
150.826 Million cell updates/sec

Title: US-10-031-158B-14

Perfect score: 58

Sequence: 1 MQMPPSPPLPFLQLTKQSS.....RYIGKKRRATRFMDPRRGTTP 58

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size: 0

Total number of hits satisfying chosen parameters: 645

Minimum DB seq length: 8  
Maximum DB seq length: 10

Post-processing: Listing first 45 summaries

Database: PIR 80:\*

1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4	6.9	8	2	PC4131
2	4	6.9	10	2	A56633
3	4	6.9	10	2	A32543
4	3	5.2	8	2	A4960
5	3	5.2	8	2	A61348
6	3	5.2	8	2	S08995
7	3	5.2	8	2	A49823
8	3	5.2	8	2	A28004
9	3	5.2	8	2	A43976
10	3	5.2	8	2	B43976
11	3	5.2	8	2	T10077
12	3	5.2	8	2	S19288
13	3	5.2	8	2	S21288
14	3	5.2	8	2	A05169
15	3	5.2	8	2	JS0316
16	3	5.2	8	2	S71919
17	3	5.2	8	2	A61102
18	3	5.2	9	2	PT0326
19	3	5.2	9	2	S19523
20	3	5.2	9	2	A31576
21	3	5.2	9	2	PC7076
22	3	5.2	10	2	A60624
23	3	5.2	10	2	A60410
24	3	5.2	10	2	S08997
25	3	5.2	10	2	A60421
26	3	5.2	10	2	S08998
27	3	5.2	10	2	A26381
28	3	5.2	10	2	JN0440
29	3	5.2	10	2	P00753

30	3	5.2	10	2	A61622	vitellogenin, 190k
31	3	5.2	10	2	C39111	Ig heavy chain C r
32	3	5.2	10	2	S10785	enamelin, 22K - bo
33	3	5.2	10	2	S70722	65.4K GTP-binding
34	3	5.2	10	2	PT0084	protein QA600021 -
35	3	5.2	10	2	B59272	peptide-N4-(N-acet
36	3	5.2	10	2	A43977	transcription fact
37	3	5.2	10	2	A42089	angiotensin precu
38	3	5.2	10	2	A90917	cytochrome-c oxida
39	3	5.2	10	2	T17075	neuropeptide Pec-H
40	3	5.2	10	2	S53789	NADH2 dehydrogenas
41	3	5.2	10	2	PQ0784	sperm-activating p
42	3	5.2	10	2	G60787	sperm-activating p
43	3	5.2	10	2	G60787	sperm-activating p
44	3	5.2	10	2	G60787	sperm-activating p
45	3	5.2	10	2	C60588	sperm-activating p

#### ALIGNMENTS

##### RESULT 1

PC4131 hypothetical protein 8 [imported] - Pseudomonas aeruginosa (fragment)

C:Species: Pseudomonas aeruginosa

C>Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 31-Dec-2004

C:Accession: PC4131

R:Kawasaki, S.; Arai, H.; Igarashi, Y.; Kodama, T.

Gene 167, 87-91, 1995

A:Title: Sequencing and characterization of the downstream region of the genes encoding

Y for biosynthesis of heme d1.

A:Reference number: J04552; PMID:96144254; PMID:8566817

A:Accession: PC4131

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-8 <RAW>

A:References: UNIPROT:P95412; UNIPARC:UPI000017A96A; DDBJ:D50473; NID:g1217594

A>Note: this ORF is not annotated in GenBank entry PSEWIRC, release 113.0

C:Superfamily: Pseudomonas stutzeri nird protein

Query Match 6.9%; Score 4; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 SRRL 23

Db 5 SRRL 8

##### RESULT 2

A56633 neomyosuppressin - flesh fly (Sarcophaga bullata)

N:Alternate names: Neb-MS

C:Species: Sarcophaga bullata

C>Date: 21-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 09-Jul-2004

C:Accession: A56633

R:Fontagy, A.; Schoofs, L.; Proost, P.; Van Damme, J.; Bueds, H.; De Loof, A.

Comp. Biochem. Physiol. C 102, 239-245, 1992

A:Title: Isolation, primary structure and synthesis of neomyosuppressin, a myoinhibiting

A:Reference number: A56633; PMID:93047886; PMID:1358537

A:Accession: A56633

A:Molecule type: protein

A:Residues: 1-10 <FON>

A:Cross-references: UNIPROT:P61850; UNIPARC:UPI000003AD0A

A:Experimental source: head

A>Note: sequence extracted from NCBI backbone (NCBIP:119072)

C:Keywords: amidated carboxyl end; neuropeptide

F:10/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 6.9%; Score 4; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 VFLLR 31  
|||  
Db 6 VFLLR 9

## RESULT 3

A:Accession: A32543  
cardioexcitatory neuropeptide - desert locust  
C:Species: Schistocerca gregaria (desert locust)  
C:Date: 20-Dec-1989 #sequence\_revision 20-Dec-1989 #text\_change 05-Oct-2004  
A:Accession: A32543  
R:Robb, S.; Packman, L.C.; Evans, P.D.  
Biochem. Biophys. Res. Commun. 160, 850-856, 1989  
A:Title: Isolation, primary structure and bioactivity of Schistocerca gregaria, a FMRF-amide  
A:Reference number: A32543; MUID:89246543; PMID:2719702  
A:Accession: A32543  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-10 <ROB>  
A:Cross-references: UNIPROT:P38553; UNIPARC:UPI000003AD64  
C:Keywords: amidated carboxyl end; neuropeptide  
F:1/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 6.9%; Score 4; DB 2; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1e+03;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 VFLLR 31  
|||  
Db 6 VFLLR 9

## RESULT 4

A:Accession: A44960  
neuropeptide Iad-CC-I - Colorado potato beetle  
C:Species: Leptinotarsa decemlineata (Colorado potato beetle)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
A:Accession: A44960  
R:Gaede, G.; Kellner, R.  
Peptides 10, 1287-1289, 1989  
A:Title: The metabolic neuropeptides of the corpus cardiacum from the potato beetle and  
A:Reference number: A44960; MUID:90160053; PMID:2576128  
A:Accession: A44960  
A:Molecule type: protein  
A:Residues: 1-8 <GAB>  
A:Cross-references: UNIPROT:P04548; UNIPARC:UPI000012CDBB  
C:Superfamily: adipokinetic hormone  
C:Keywords: blocked carboxyl end; corpora cardiaca; hormone; neuropeptide; pyroglutamic  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:8/Modified site: blocked carboxyl end (Trp) (probably amidated) #status experimental

Query Match 5.2%; Score 3; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 NFS 34  
|||  
Db 3 NFS 5

## RESULT 5

A:Accession: A61348  
red pigment-concentrating hormone - northern shrimp  
N:Alternate names: blanching hormone  
C:Species: Penaeus borealis (northern shrimp)  
C:Date: 02-Aug-1994 #sequence\_revision 05-Aug-1994 #text\_change 09-Jul-2004  
A:Accession: A61348; S07139  
R:Fernlund, P.; Josefsson, L.  
Science 177, 173-175, 1972  
A:Title: Crustacean color-change hormone: amino acid sequence and chemical synthesis.  
A:Reference number: A61348; MUID:72228738; PMID:5041363  
A:Accession: A61348  
A:Molecule type: protein

A:Residues: 1-8 <FER1>  
A:Cross-references: UNIPROT:P08939; UNIPARC:UPI000013460C  
R:Fernlund, P.  
Biochim. Biophys. Acta 371, 304-311, 1974  
A:Title: Structure of the red-pigment-concentrating hormone of the shrimp, Penaeus borealis  
A:Reference number: S07139; MUID:75054965; PMID:4433569  
A:Accession: S07139  
A:Molecule type: protein  
A:Residues: 'E', 2-8 <FER2>  
A:Cross-references: UNIPARC:UPI000017661E  
A:Note: the amino-terminal residue forms pyrrolidone carboxylic acid; therefore, we have  
zed pigment-containing cells.

C:Superfamily: adipokinetic hormone  
C:Keywords: amidated carboxyl end; blocked amino end; hormone; neuropeptide; pyroglutamic  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:8/Modified site: amidated carboxyl end (Trp) #status experimental

Query Match 5.2%; Score 3; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 NFS 34  
|||  
Db 3 NFS 5

## RESULT 6

S08995  
hypertrehalosemic hormone I - oriental cockroach  
N:Alternate names: Pea-CAH-I  
C:Species: Blatta orientalis (oriental cockroach)  
C:Date: 30-Jun-1992 #sequence\_revision 24-Oct-1997 #text\_change 09-Jul-2004  
A:Accession: S08995  
R:Gaede, G.; Rinehart, K.L.  
Biol. Chem. Hoppe-Seyler 371, 345-354, 1990  
A:Title: Primary structures of hypertrehalosemic neuropeptides isolated from the corpora  
entails and of the stick insect Ectatostoma tlaxiatum assigned by tandem fast atom bombard

A:Reference number: S08995; MUID:90253659; PMID:2340112  
A:Accession: S08995  
A:Molecule type: protein  
A:Residues: 1-8 <GAB>  
A:Cross-references: UNIPROT:P04548; UNIPARC:UPI000012CDBB  
A:Note: the amino-terminal residue forms pyrrolidone carboxylic acid; therefore, we have  
C:Superfamily: adipokinetic hormone  
C:Keywords: amidated carboxyl end; hormone; neuropeptide; pyroglutamic acid  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:8/Modified site: amidated carboxyl end (Trp) #status experimental

Query Match 5.2%; Score 3; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 NFS 34  
|||  
Db 3 NFS 5

## RESULT 7

A:Accession: A49823  
adipokinetic hormone I - American cockroach  
N:Alternate names: periplanetin CC-1  
C:Species: Periplaneta americana (American cockroach)  
C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 09-Jul-2004  
A:Accession: A49823  
R:Scarborough, R.M.; Jamieson, G.C.; Kallish, F.; Krmer, S.J.; McBurnee, G.A.; Miller, C.  
Proc. Natl. Acad. Sci. U.S.A. 81, 5575-5579, 1984  
A:Title: Isolation and primary structure of two peptides with cardioacceleratory and hyp  
A:Reference number: A49823; MUID:84298179; PMID:6591205  
A:Accession: A49823  
A:Molecule type: protein  
A:Residues: 1-8 <SCA>  
A:Cross-references: UNIPROT:P04548; UNIPARC:UPI000012CDBB

C:Superfamily: adipoiknetic hormone  
C:Keywords: amidated carboxyl end; corpora cardiaca; hormone; neuropeptide; pyroglutamic  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:8/Modified site: amidated carboxyl end (Tyr) #status experimental

Query Match 5.2%; Score 3; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 32 NFS 34  
|||  
Db 3 NFS 5

RESULT 8  
A28004  
adipoiknetic hormone G - two-spotted cricket  
N:Alternate names: AKH-G

C:Species: Gryllus bimaculatus (two-spotted cricket)  
C>Date: 30-Jun-1989 #sequence\_revision 24-Oct-1997 #text\_change 09-Jul-2004  
C:Accession: A28004

R:Gaede, G.; Rinehart, K.L.  
Biochem. Biophys. Res. Commun. 149, 908-914, 1987  
A:Title: Primary sequence analysis by fast atom bombardment mass spectrometry of a peptide  
A:Reference number: A28004; PMID:8810653; PMID:3426616  
A:Accession: A28004

A:Molecule type: Protein  
A:Residues: 1-8 <GAE>

A:Cross-references: UNIPROT:P14086; UNIPARC:UPI000012576F

A>Note: the amino-terminal residue forms pyrrolidone carboxylic acid; therefore, we have  
C:Superfamily: adipoiknetic hormone

C:Keywords: amidated carboxyl end; corpora cardiaca; hormone; neuropeptide; pyroglutamic  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:8/Modified site: amidated carboxyl end (Tyr) #status experimental

Query Match 5.2%; Score 3; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 32 NFS 34  
|||  
Db 3 NFS 5

RESULT 9  
A43976

hypertrehalosemic hormone - yellow mealworm

C:Species: Tenebrio molitor (yellow mealworm)

C>Date: 03-Feb-1993 #sequence\_revision 03-Feb-1993 #text\_change 09-Jul-2004  
C:Accession: A43976

R:Gaede, G.; Rosinski, G.  
Peptides 11, 455-459, 1990

A:Title: The primary structure of the hypertrehalosemic neuropeptide from tenebrionid be  
A:Reference number: A43976; PMID:90341081; PMID:2381871  
A:Accession: A43976

A:Molecule type: Protein  
A:Residues: 1-8 <GAE>

A:Cross-references: UNIPROT:P25419; UNIPARC:UPI000012CDC8

C:Superfamily: adipoiknetic hormone

C:Keywords: amidated carboxyl end; corpora cardiaca (Gln) #status experimental  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:8/Modified site: amidated carboxyl end (Tyr) #status experimental

Query Match 5.2%; Score 3; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 32 NFS 34  
|||  
Db 3 NFS 5

RESULT 10

B43976  
hypertrehalosemic hormone - beetle (Zophobas rugipes)

C:Species: Zophobas rugipes  
C>Date: 03-Feb-1993 #sequence\_revision 03-Feb-1993 #text\_change 09-Jul-2004  
C:Accession: B43976

R:Gaede, G.; Rosinski, G.  
Peptides 11, 455-459, 1990

A:Title: The primary structure of the hypertrehalosemic neuropeptide from tenebrionid be  
A:Reference number: A43976; PMID:90341081; PMID:2381871  
A:Accession: B43976

A:Molecule type: Protein  
A:Residues: 1-8 <GAE>

A:Cross-references: UNIPROT:P25419; UNIPARC:UPI000012CDC8

C:Superfamily: adipoiknetic hormone

C:Keywords: amidated carboxyl end; corpora cardiaca; hormone; neuropeptide; pyroglutamic  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:8/Modified site: amidated carboxyl end (Tyr) #status experimental

Query Match 5.2%; Score 3; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 32 NFS 34  
|||  
Db 3 NFS 5

RESULT 11  
T10077

hypothetical protein N - Methylophilus methylotrophus (fragment)

C:Species: Methylophilus methylotrophus

C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
C:Accession: T10077

R:Christoserdov, A.V.; McIntire, W.S.; Mathews, F.S.; Lidstrom, M.E.  
J. Bacteriol. 176, 4073-4080, 1994

A:Title: Organization of the methylamine utilization (mau) genes in Methylophilus methyl  
A:Reference number: Z16936; PMID:94292427; PMID:8021188  
A:Accession: T10077

A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA

A:Residues: 1-8 <CHI>

A:Cross-references: UNIPARC:UPI000011E8A1; EMBL:L26407; NID:G561931; PIDN:AAB46955.1; PII  
C:Genetics: mauN

Query Match 5.2%; Score 3; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 21 RRL 23  
|||  
Db 2 RRL 4

RESULT 12  
S19288

acylase - Kluyvera cryocrescens

C:Species: Kluyvera cryocrescens

C>Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
C:Accession: S19288

R:Martin, U.; Slade, A.; Aitken, A.; Arche, R.; Vidén, R.  
Biochem. J. 280, 659-662, 1991

A:Title: Chemical modification of serine at the active site of penicillin acylase from KJ  
A:Reference number: S19288; PMID:1764029  
A:Accession: S19288

A:Status: preliminary  
A:Molecule type: Protein

A:Residues: 1-8 <MAK>

A:Cross-references: UNIPROT:Q7M124; UNIPARC:UPI000017AA77

Query Match 5.2%; Score 3; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.8e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 41 IGX 43  
 |||  
 Db 6 IGX 8

## RESULT 13

S21288  
 lectin - potato (fragment)  
 C:Species: Solanum tuberosum (potato)  
 C:Date: 19-Mar-1997 #sequence\_revision 05-Dec-1998 #text\_change 09-Jul-2004  
 C:Accession: S21288  
 R:Miller, D.J.; Allen, A.K.; Smith, C.G.; Sidebottom, C.; Slabas, A.R.; Bolwell, G.P.  
 Biochem. J. 283, 813-821, 1992  
 A:Title: Chitin-binding proteins in potato (Solanum tuberosum L.) tuber. Characterization  
 A:Reference number: S21288; MUID:92272683; PMID:1590771  
 A:Accession: S21288  
 A:Molecule type: protein  
 A:Residues: 1-8 <MIT>  
 A:Cross-references: UNIPROT:Q7M1V6; UNIPARC:UPI000017B0BF  
 A:Experimental source: var. Ulster Sceptre  
 C:Function:  
 A:Description: may be involved in defence mechanism of the plant  
 C:Keywords: hydroxyproline; lectin

Query Match 5.2%; Score 3; DB 2; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 PSP 8  
 |||  
 Db 5 PSP 7

## RESULT 14

A05169  
 neuropeptide M-I - American cockroach  
 C:Species: Periplaneta americana (American cockroach)  
 C:Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 09-Jul-2004  
 C:Accession: A05169  
 R:Witten, J.L.; Schaffer, M.H.; O'Shea, M.; Cook, J.C.; Hemling, M.E.; Rinehart Jr., K.L.  
 Biochem. Biophys. Res. Commun. 124, 350-358, 1984  
 A:Title: Structures of two cockroach neuropeptides assigned by fast atom bombardment mass  
 A:Reference number: A05118; MUID:85046530; PMID:6548628  
 A:Accession: A05169  
 A:Molecule type: protein  
 A:Residues: 1-8 <MIT>  
 A:Cross-references: UNIPROT:P04548; UNIPARC:UPI000017BE15  
 C:Keywords: neuropeptide

Query Match 5.2%; Score 3; DB 2; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 NFS 34  
 |||  
 Db 3 NFS 5

## RESULT 15

US0316  
 leucokinin VI - Madeira cockroach  
 C:Species: Leucophaea maderae (Madeira cockroach)  
 C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 09-Jul-2004  
 C:Accession: US0316  
 R:Holman, G.M.; Cook, B.J.; Nachman, R.J.  
 Comp. Biochem. Physiol. C 89, 27-30, 1987  
 A:Title: Isolation, primary structure, and synthesis of leucokinin V and VI: myotropic  
 A:Reference number: US0315  
 A:Accession: US0316  
 A:Molecule type: protein  
 A:Residues: 1-8 <HOL>

A:Cross-references: UNIPROT:P19988; UNIPARC:UPI000012E29E  
 C:Comment: Leucokinin, a family of cephalomyotropic peptides, stimulate contractile act  
 C:Keywords: amidated carboxyl end; cephalomyotropic peptide; pyroglutamic acid  
 F:1/Modified site: pyroglutamic acid (Gln) #status experimental  
 F:8/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match 5.2%; Score 3; DB 2; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 QSS 20  
 |||  
 Db 1 QSS 3

Search completed: December 13, 2005, 08:27:10  
 Job time : 39 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2005, 08:18:33 ; Search time 226 Seconds

(without alignments)  
181.065 Million cell updates/sec

Title: US-10-031-158b-14

Perfect score: 58

Sequence: 1 MQMFPSPPLFFFLQLKQSS.....RYIGKKRRATRFMDPRRGTG 58

Scoring table: OLIGO

Searched: 2166443 seqs, 705528306 residues

Word size: 0

Total number of hits satisfying chosen parameters: 2877

Minimum DB seq length: 8

Maximum DB seq length: 10

Post-processing: Listing first 45 summaries

Database: UniProt 05.80:\*

1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4	6.9	8	015896_BABBO	015896 babesia boy
2	4	6.9	9	088889_MOUSE	088889 mus musculu
3	4	6.9	10	FARP_LOCM1	P84306 locusta mlg
4	4	6.9	10	FARP_SCHGR	P84307 schistocerc
5	4	6.9	10	LCMS_LEUMA	P11444 leucophaea
6	4	6.9	10	NEMS_SARBU	P61850 sarcophaga
7	4	6.9	10	06A3T0_9FILI	06a3t0 archangiopt
8	4	6.9	10	Q76M18_9TELE	Q76m18 eurypathryn
9	3	5.2	8	AKHG_GRYBI	P67785 gryllus bim
10	3	5.2	8	AKH_ROMMI	P67786 romalea mic
11	3	5.2	8	CADI_ENTFA	P13268 enterococcu
12	3	5.2	8	FAR1_PANRE	P41872 panagrellus
13	3	5.2	8	FAR1_PENMO	P83316 penaeus mon
14	3	5.2	8	FAR2_MACRS	P83275 macrodrachi
15	3	5.2	8	FAR3_HOMAM	P41486 homartus ame
16	3	5.2	8	FAR4_HOMAM	P41487 homartus ame
17	3	5.2	8	HFP1_BLAOR	P84261 blattella orie
18	3	5.2	8	HFP1_LEPDE	P84260 lepidinotars
19	3	5.2	8	HFP1_PERAM	P84259 periplaneta
20	3	5.2	8	HFP_TENNO	P67788 tenebrio mo
21	3	5.2	8	HFP_ZOPRU	P67790 zophobas ru
22	3	5.2	8	LC6K_LEUMA	P19988 leucophaea
23	3	5.2	8	RPCH_PANBO	P08939 pandanus bo
24	3	5.2	8	013591_YEAST	013591 saccharomyc
25	3	5.2	8	P82858_9BASI	P82858 puccinia re
26	3	5.2	8	Q69YH8_HUMAN	Q69yh8 homo sapien
27	3	5.2	8	Q6TF70_HUMAN	Q6tf70 homo sapien
28	3	5.2	8	Q75MD1_HUMAN	Q75md1 homo sapien
29	3	5.2	8	Q68B59_STRPU	Q68b59 strongyloce
30	3	5.2	8	Q37854_BPR17	Q37854 bacterioph
31	3	5.2	8	Q40530_TOBAC	Q40530 nicotiana t

32	3	5.2	8	06JC68_SOYAN	06jc68 glycine max
33	3	5.2	8	070Y84_9LAMI	070y84 plectranthu
34	3	5.2	8	07MIV6_SOLMU	07miv6 solanum tub
35	3	5.2	8	Q9RS80_SHIDY	Q9rs80 shigella dy
36	3	5.2	8	Q79F63_SYNSP	Q79f63 synecococc
37	3	5.2	8	Q7M124_KLUCI	Q7m124 kluyvera ci
38	3	5.2	8	Q78DX6_RAT	Q78dx6 rattus norv
39	3	5.2	8	08R5M9_MOUSE	08r5m9 mus musculu
40	3	5.2	8	Q9ET16_MESAU	Q9et16 mesocricetu
41	3	5.2	8	Q9ET17_MUSCR	Q9et17 mus caroli
42	3	5.2	8	Q9ET18_MUSSP	Q9et18 mus spretus
43	3	5.2	8	Q6PUD5_SV40	Q6pud5 simian viru
44	3	5.2	8	Q6PUD7_SV40	Q6pud7 simian viru
45	3	5.2	8	Q6PUD9_SV40	Q6pud9 simian viru

#### ALIGNMENTS

RESULT 1

ID	015896_BABBO	PRELIMINARY;	PRT;	8 AA.
AC	015896;			
DT	01-JAN-1998 (TREMBlrel. 05, Created)			
DT	01-JAN-1998 (TREMBlrel. 05, Last sequence update)			
DT	01-FEB-2005 (TREMBlrel. 29, Last annotation update)			
DE	12D3 antigen (Fragment).			
GN	Name=12D3;			
OS	Babesia bovis.			
OC	Eukaryota; Alveolata; Apicomplexa; Piroplasmida; Babesiidae; Babesia.			
OX	NCBI_TaxID=5865;			
RN	[1]			
RP	NCLEBOTIDE SEQUENCE.			
RC	STRAIN=Samford attenuated;			
RA	Silene G.U., Blakeley R.L., Riddles P.W.;			
RL	Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	NCLEBOTIDE SEQUENCE.			
RA	Silene G., Blakeley R., Riddles P.;			
RL	Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; U4917; BAB6362.1; -; Genomic DNA.			
DR	EMBL; AY170919; AAN64587.1; -; Genomic DNA.			
FT	NON TER			
SQ	SEQUENCE 8 AA; 984 MW; FDD2C9D411ADD726 CRC64;			

Query Match 6.9%; Score 4; DB 2; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.2e+06;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 ATRF 50  
Db 3 ATRF 6

RESULT 2

ID	088889_MOUSE	PRELIMINARY;	PRT;	9 AA.
AC	088889;			
DT	01-NOV-1998 (TREMBlrel. 08, Created)			
DT	01-NOV-1998 (TREMBlrel. 08, Last sequence update)			
DT	01-MAR-2004 (TREMBlrel. 26, Last annotation update)			
DE	Ubiquitin-conjugating enzyme UBCH4 (Fragment).			
GN	Name=Ube213; Synonym=UBCH4;			
OS	Mus musculus (mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;			
OC	Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	NCLEBOTIDE SEQUENCE.			
RC	STRAIN=129/Sv;			
RX	MEDLINE=7057256; PubMed=8901595; DOI=10.1073/pnas.93.22.12412;			
RA	Harbers K., Mueller U., Grams A., Li E., Jaenisch R., Franz T.;			

RT "Provirus integration into a gene encoding a ubiquitin-conjugating  
 RT enzyme results in a placental defect and embryonic lethality."  
 RL Proc. Natl. Acad. Sci. U.S.A. 93:12412-12417(1996).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=128/Sv;  
 RA MEDLINE=99132641;  
 RM Mller U., Grams A., Martinez-Noel G., Copeland N.G., Gilbert D.J.,  
 RA Jenkins N.A., Harbers K.;  
 RT "Structure of the gene encoding the ubiquitin-conjugating enzyme  
 RT UbcM4, characterization of its promoter, and chromosomal location."  
 RL Gene 224:109-116(1998).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=129/Sv;  
 RA Mueller U., Grams A., Martinez-Noel G., Harbers K.;  
 RL Submitted (JUN-1998) to the EMBL/Genbank/DBJ databases.  
 DR EMBL, AF071557; AAD10126.1; -, Genomic\_DNA.  
 DR MGI; MGI:109240; Ube213.  
 FT NON TER  
 SQ SEQUENCE 9 AA; 1063 MW; C90F97341415BDD CRC64;  
 Query Match 6.9%; Score 4; DB 2; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+06; Mismatches 0; Indels 0; Gaps 0;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 20 SRRL 23  
 Db 4 SRRL 7  
 RESULT 3  
 FARP LOCMI STANDARD; PRT; 10 AA.  
 AC P84307; P38553;  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 01-FEB-2005 (Rel. 46, Last annotation update)  
 DE Schistocerca gregaria (Schistocerca gregaria) (Cardioexcitatory neuropeptide).  
 OS Locusta migratoria (Migratory locust).  
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 CC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridoidea;  
 CC Acrididae; Acrididae; Oedipodinae; Locusta.  
 CC NCBI\_TaxID=7004;  
 RN [1]  
 RP PROTEIN SEQUENCE, FUNCTION, SUBCELLULAR LOCATION, AND TISSUE  
 RP SPECIFICITY.  
 RC TISSUE=Corpora cardiaca;  
 RX MEDLINE=93324430; PubMed=7687352; DOI=10.1016/0196-9781(93)90126-2;  
 RA Schoofs L., Holman G.M., Paemen L., Veelaert D., Amelincx M.,  
 RA de Loof A.;  
 RT "Isolation, identification, and synthesis of PDVDFLRFamide  
 RT (Schistocerca gregaria) in Locusta migratoria and its association with the  
 RT male accessory glands, the salivary glands, the heart, and the  
 RT oviduct."  
 RL Peptides 14:409-421(1993).  
 CC -1- FUNCTION: Muscle inhibiting agent. Involved in the neural control  
 CC of the visceral muscles of the heart, accessory glands and  
 CC oviduct. May be involved in the regulation of saliva secretion.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: Found in axons of the male accessory glands,  
 CC the salivary glands, the heart, and the oviduct.  
 CC -1- SIMILARITY: Belongs to the FARP (FMRamide related peptide)  
 CC family.  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC Amidation: Direct protein sequencing; Neuropeptide.  
 CC MOD\_RES 10 10 Phenylalanine amide.  
 FT

SQ SEQUENCE 10 AA; 1244 MW; D3C51729D2C1EAB2 CRC64;  
 Query Match 6.9%; Score 4; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 6.2e+03; Mismatches 0; Indels 0; Gaps 0;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 28 VFRL 31  
 Db 6 VFRL 9  
 RESULT 4  
 FARP SCHGR STANDARD; PRT; 10 AA.  
 AC P84307; P38553;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 01-FEB-2005 (Rel. 46, Last annotation update)  
 DE Schistocerca gregaria (Desert locust).  
 OS Schistocerca gregaria (Desert locust).  
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 CC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridoidea;  
 CC Acrididae; Acrididae; Cyrtacanthacridinae; Schistocerca.  
 CC NCBI\_TaxID=7010;  
 RN [1]  
 RP PROTEIN SEQUENCE, AND FUNCTION.  
 RP TISSUE=Thoracic nervous system;  
 RX MEDLINE=89246543; PubMed=2719702;  
 RA Robb S., Packman L.C., Evans P.D.;  
 RT "Isolation, primary structure and bioactivity of schistocerca gregaria, a  
 RT FMRF-amide-like neuropeptide from the locust, Schistocerca gregaria."  
 RL Biochem. Biophys. Res. Commun. 160:850-856(1989).  
 CC -1- FUNCTION: Muscle inhibiting agent. Involved in the neural control  
 CC of the visceral muscles of the heart, accessory glands and  
 CC oviduct. May be involved in the regulation of saliva secretion.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: Belongs to the FARP (FMRamide related peptide)  
 CC family.  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC PIR: A32543; A32543.  
 DR Amidation: Direct protein sequencing; Neuropeptide.  
 FT MOD\_RES 10 10 Phenylalanine amide.  
 SQ SEQUENCE 10 AA; 1244 MW; D3C51729D2C1EAB2 CRC64;  
 Query Match 6.9%; Score 4; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 6.2e+03; Mismatches 0; Indels 0; Gaps 0;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 28 VFRL 31  
 Db 6 VFRL 9  
 RESULT 5  
 LCM5\_LBDUA STANDARD; PRT; 10 AA.  
 ID LCM5\_LBDUA  
 AC P21144; P41497;  
 DT 01-MAY-1991 (Rel. 18, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Leucomyosuppressin (LMS) (Lem-MS).  
 OS Leucomyosuppressin (LMS) (Lem-MS).  
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 CC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blaberidae;  
 CC Blaberidae; Leucomyosuppressin.  
 CC NCBI\_TaxID=6988;  
 RN [1]

RP PROTEIN SEQUENCE, AND SYNTHESIS.  
 RC TISSUE=Head;  
 RA Holman G.M., Cook B.J., Nachman R.J.;  
 RT "Isolation, primary structure and synthesis of leucenyosuppressin, an  
 insect neuropeptide that inhibits spontaneous contractions of the  
 cockroach hindgut";  
 RL Somp. Biochem. Physiol. 85C:329-333(1986).  
 CC -1- FUNCTION: Inhibits the spontaneous contractions of cockroach  
 proctodum (hindgut).  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use as long as its content is in no way modified and this statement is not  
 removed.  
 CC -----  
 KW Amidation; Direct protein sequencing; Neuropeptide;  
 KM Pyrrolidone carboxylic acid.  
 FT MOD\_RES 1 1 Pyrrolidone carboxylic acid.  
 FT MOD\_RES 10 10 Phenylalanine amide.  
 SQ SEQUENCE 10 AA; 1275 MW; D3C45229D2C1EAB2 CRC64;  
 Query Match 6.9%; Score 4; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 6.2e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 28 VFLLR 31  
 Db 6 VFLLR 9

RESULT 6  
 NEMS\_SARBU STANDARD; PRT; 10 AA.  
 ID NEMS\_SARBU  
 AC P61850; P41494; Q9VC91;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 05-JUL-2004 (Rel. 44, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Neomysuppressin (Neb-MS) (TVDVHVFLLRamide).  
 GN Name=NEMS;  
 OS Sarcophaga bullata (Grey flesh fly) (Neobellieria bullata).  
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;  
 CC Sarcophagidae; Sarcophaga; Neobellieria.  
 CC NCBL\_TaxID=7385;  
 RN [1]  
 RP PROTEIN SEQUENCE.  
 RC TISSUE=Head;  
 RX MEDLINE=93047886; PubMed=1358537;  
 RA Fonagy A., Schoofs L., Proost P., Van Damme J., Buede H., De Loof A.;  
 RT "Isolation, primary structure and synthesis of neomysuppressin, a  
 myoinhibiting neuropeptide from the grey fleshfly, Neobellieria  
 bullata";  
 RL Comp. Biochem. Physiol. 102C:239-245(1992).  
 CC -1- FUNCTION: Myoinhibiting neuropeptide.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use as long as its content is in no way modified and this statement is not  
 removed.  
 CC -----  
 KW PIR; A56633; A56633.  
 KM Amidation; Direct protein sequencing; Neuropeptide.  
 FT MOD\_RES 10 10 Phenylalanine amide.  
 SQ SEQUENCE 10 AA; 1248 MW; D3C00329D2C1EAB2 CRC64;  
 Query Match 6.9%; Score 4; DB 1; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 6.2e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 VFLLR 31  
 Db 6 VFLLR 9

RESULT 7  
 ID 06A3T0\_9FILLI PRELIMINARY; PRT; 10 AA.  
 AC 06A3T0;  
 DT 25-OCT-2004 (TREMBlrel. 28, Created)  
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)  
 DE ATP synthase beta subunit (EC 3.6.3.14) (Fragment).  
 GN Name=atp3;  
 OS Archangiopteris somai.  
 CC Chloroplast.  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Monilliformes; Filicophyta; Marattiales; Marattiales;  
 CC Marattiaceae; Archangiopteris.  
 CC NCBL\_TaxID=203826;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Chiang T., Chiang Y., Chou C., Cheng Y., Chiou W.;  
 RT "Phylogeny and conservation of Archangiopteris somai and A. itoi  
 (Marattiales, Pteridophyta) based on nucleotide variation of cpDNA  
 atp3-rbcL intergenic spacer";  
 RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Chiang Y.C.;  
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ505259; CAD44048.1; -; Genomic\_DNA.  
 DR GO; GO:0009507; C:chloroplast; IEA.  
 DR GO; GO:0016787; F:hydrolyase activity; IEA.  
 KW Chloroplast; Hydrolase.  
 KM NON\_TER 10 10  
 SQ SEQUENCE 10 AA; 1195 MW; 9E2AF0C9C7376451 CRC64;  
 Query Match 6.9%; Score 4; DB 2; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 6.2e+03;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 8 PLFF 11  
 Db 5 PLFF 8

RESULT 8  
 ID 076ML8\_9TELE PRELIMINARY; PRT; 10 AA.  
 AC 076ML8;  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DT 01-FEB-2005 (TREMBlrel. 29, Last annotation update)  
 DE NMDH dehydrogenase subunit 6 (Fragment).  
 GN Name=ND6;  
 OS Burypharynx pelicanoides (pelican eel).  
 CC Mitochondrion.  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Actinopterygii; Neopterygii; Teleostei; Anguilliformes;  
 CC Eurypharyngidae; Burypharynx.  
 CC NCBL\_TaxID=55117;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=22967687; PubMed=12949142; DOI=10.1093/molbev/meg206;  
 RA Inoue J.G., Miya M., Tsukamoto K., Nishida M.;  
 RT "Evolution of the deep-sea gulper eel mitochondrial genomes: large-  
 scale gene rearrangements originated within the eels";  
 RL Mol. Biol. Evol. 20:1917-1924(2003).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Inoue J., Inoue J., Miya M., Tsukamoto K., Nishida M.;  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

```
DR EMBL; AB046481; BAB87147.1; -; Genomic_DNA.
DR EMBL; AB046485; BAB87155.1; -; Genomic_DNA.
DR EMBL; AB046489; BAB87163.1; -; Genomic_DNA.
DR EMBL; AB046477; BAB87139.1; -; Genomic_DNA.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 10 10
SQ SEQUENCE 10 AA; 1297 MW; D43ECB9C9C059C9 CRC64;

Query Match
Best Local Similarity 100.0%; Score 4; DB 2; Length 10;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 33 FSLM 36
Db 7 FSLM 10

RESULT 9
AKHG_GRYBI STANDARD; PRT; 8 AA.
ID _AKHG_GRYBI
AC P67785; P14086;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Adipokinetic hormone G (AKH-G).
OS Gryllus bimaculatus (Two-spotted cricket).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthoptera; Orthoptera; Ensifera; Gryllidae; Gryllinae;
OC Gryllus.
OC NCBI_TaxID=6999;
RN [1]
RP PROTEIN SEQUENCE.
RC TISSUE=Corpora cardiaca;
RX MEDLINE=88106553; PubMed=3426616;
RA Gaede G., Rinehart K.L. Jr.;
RT "Primary sequence analysis by fast atom bombardment mass spectrometry
RT of a peptide with adipokinetic activity from the corpora cardiaca of
RT the cricket Gryllus bimaculatus."
RL Biochem. Biophys. Res. Commun. 149:908-914(1987).
CC -1- FUNCTION: This hormone, released from cells in the corpora
CC cardiaca after the beginning of flight, causes release of
CC diglycerides from the fat body and then stimulates the flight
CC muscles to use these diglycerides as an energy source.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the AKH/HRTM/RPCH family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC PIR: A28004; A28004.
DR InterPro; IPR002047; AKH.
DR PROSITE; PS00256; AKH; 1.
KW Amidation; Direct protein sequencing; Flight; Neuropeptide;
KW Pyroglutamate carboxylic acid.
FT MOD_RES 1 1 Pyroglutamate carboxylic acid.
FT MOD_RES 8 8 Tryptophan amide.
SQ SEQUENCE 8 AA; 938 MW; 867861B5B9C452D6 CRC64;

Query Match
Best Local Similarity 100.0%; Score 3; DB 1; Length 8;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 32 NFS 34
Db 3 NFS 5

RESULT 10
AKH_ROMMI STANDARD; PRT; 8 AA.
ID _AKH_ROMMI
AC P67786; P14086;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Adipokinetic hormone (AKH) (RO II).
OS Romalea microptera (lubber grasshopper).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthoptera; Orthoptera; Caelifera; Acridoformia;
OC Acridoidea; Romaleidae; Romalea.
OC NCBI_TaxID=7007;
RN [1]
RP PROTEIN SEQUENCE.
RC TISSUE=Corpora cardiaca;
RX MEDLINE=89145002; PubMed=3226948; DOI=10.1016/0196-9781(88)90107-6;
RA Gaede G., Hilbich C., Beyreuther K., Rinehart K.L. Jr.;
RT "Sequence analyses of two neuropeptides of the AKH/RPCH family from
RT the lubber grasshopper, Romalea microptera."
RL Peptides 9:681-688(1986).
CC -1- FUNCTION: This hormone, released from cells in the corpora
CC cardiaca after the beginning of flight, causes release of
CC diglycerides from the fat body and then stimulates the flight
CC muscles to use these diglycerides as an energy source.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the AKH/HRTM/RPCH family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR InterPro; IPR002047; AKH.
DR PROSITE; PS00256; AKH; 1.
KW Amidation; Direct protein sequencing; Flight; Neuropeptide;
KW Pyroglutamate carboxylic acid.
FT MOD_RES 1 1 Pyroglutamate carboxylic acid.
FT MOD_RES 8 8 Tryptophan amide.
SQ SEQUENCE 8 AA; 938 MW; 867861B5B9C452D6 CRC64;

Query Match
Best Local Similarity 100.0%; Score 3; DB 1; Length 8;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 32 NFS 34
Db 3 NFS 5

RESULT 11
CAD1_ENTFA STANDARD; PRT; 8 AA.
ID _CAD1_ENTFA
AC P13268;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 05-JUN-2004 (Rel. 44, Last annotation update)
DE Sex pheromone CAD1.
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OC NCBI_TaxID=151;
RN [1]
RP PROTEIN SEQUENCE.
RX MEDLINE=85051889; PubMed=6437872; DOI=10.1016/0014-5793(84)81248-X;
RA Mori M., Sagakami Y., Narita M., Isegai A., Fujino M., Kitada C.,
RA Craig R.A., Clevelle D.B., Suzuki A.;
RT "Isolation and structure of the bacterial sex pheromone, CAD1, that
RT induces plasmid transfer in Streptococcus faecalis."
RL FEBS Lett. 178:97-100(1984).
CC -1- FUNCTION: Involved in the conjugative transfer of the hemolysin
CC plasmid PAD1.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
DR EMBL; AB046481; BAB87147.1; -; Genomic_DNA.
DR EMBL; AB046485; BAB87155.1; -; Genomic_DNA.
DR EMBL; AB046489; BAB87163.1; -; Genomic_DNA.
DR EMBL; AB046477; BAB87139.1; -; Genomic_DNA.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 10 10
SQ SEQUENCE 10 AA; 1297 MW; D43ECB9C9C059C9 CRC64;

Query Match
Best Local Similarity 100.0%; Score 4; DB 2; Length 10;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 33 FSLM 36
Db 7 FSLM 10

RESULT 9
AKHG_GRYBI STANDARD; PRT; 8 AA.
ID _AKHG_GRYBI
AC P67785; P14086;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Adipokinetic hormone G (AKH-G).
OS Gryllus bimaculatus (Two-spotted cricket).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthoptera; Orthoptera; Ensifera; Gryllidae; Gryllinae;
OC Gryllus.
OC NCBI_TaxID=6999;
RN [1]
RP PROTEIN SEQUENCE.
RC TISSUE=Corpora cardiaca;
RX MEDLINE=88106553; PubMed=3426616;
RA Gaede G., Rinehart K.L. Jr.;
RT "Primary sequence analysis by fast atom bombardment mass spectrometry
RT of a peptide with adipokinetic activity from the corpora cardiaca of
RT the cricket Gryllus bimaculatus."
RL Biochem. Biophys. Res. Commun. 149:908-914(1987).
CC -1- FUNCTION: This hormone, released from cells in the corpora
CC cardiaca after the beginning of flight, causes release of
CC diglycerides from the fat body and then stimulates the flight
CC muscles to use these diglycerides as an energy source.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the AKH/HRTM/RPCH family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC PIR: A28004; A28004.
DR InterPro; IPR002047; AKH.
DR PROSITE; PS00256; AKH; 1.
KW Amidation; Direct protein sequencing; Flight; Neuropeptide;
KW Pyroglutamate carboxylic acid.
FT MOD_RES 1 1 Pyroglutamate carboxylic acid.
FT MOD_RES 8 8 Tryptophan amide.
SQ SEQUENCE 8 AA; 938 MW; 867861B5B9C452D6 CRC64;

Query Match
Best Local Similarity 100.0%; Score 3; DB 1; Length 8;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 32 NFS 34
Db 3 NFS 5

RESULT 10
AKH_ROMMI STANDARD; PRT; 8 AA.
ID _AKH_ROMMI
AC P67786; P14086;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Adipokinetic hormone (AKH) (RO II).
OS Romalea microptera (lubber grasshopper).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthoptera; Orthoptera; Caelifera; Acridoformia;
OC Acridoidea; Romaleidae; Romalea.
OC NCBI_TaxID=7007;
RN [1]
RP PROTEIN SEQUENCE.
RC TISSUE=Corpora cardiaca;
RX MEDLINE=89145002; PubMed=3226948; DOI=10.1016/0196-9781(88)90107-6;
RA Gaede G., Hilbich C., Beyreuther K., Rinehart K.L. Jr.;
RT "Sequence analyses of two neuropeptides of the AKH/RPCH family from
RT the lubber grasshopper, Romalea microptera."
RL Peptides 9:681-688(1986).
CC -1- FUNCTION: This hormone, released from cells in the corpora
CC cardiaca after the beginning of flight, causes release of
CC diglycerides from the fat body and then stimulates the flight
CC muscles to use these diglycerides as an energy source.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the AKH/HRTM/RPCH family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR InterPro; IPR002047; AKH.
DR PROSITE; PS00256; AKH; 1.
KW Amidation; Direct protein sequencing; Flight; Neuropeptide;
KW Pyroglutamate carboxylic acid.
FT MOD_RES 1 1 Pyroglutamate carboxylic acid.
FT MOD_RES 8 8 Tryptophan amide.
SQ SEQUENCE 8 AA; 938 MW; 867861B5B9C452D6 CRC64;

Query Match
Best Local Similarity 100.0%; Score 3; DB 1; Length 8;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 32 NFS 34
Db 3 NFS 5

RESULT 11
CAD1_ENTFA STANDARD; PRT; 8 AA.
ID _CAD1_ENTFA
AC P13268;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 05-JUN-2004 (Rel. 44, Last annotation update)
DE Sex pheromone CAD1.
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OC NCBI_TaxID=151;
RN [1]
RP PROTEIN SEQUENCE.
RX MEDLINE=85051889; PubMed=6437872; DOI=10.1016/0014-5793(84)81248-X;
RA Mori M., Sagakami Y., Narita M., Isegai A., Fujino M., Kitada C.,
RA Craig R.A., Clevelle D.B., Suzuki A.;
RT "Isolation and structure of the bacterial sex pheromone, CAD1, that
RT induces plasmid transfer in Streptococcus faecalis."
RL FEBS Lett. 178:97-100(1984).
CC -1- FUNCTION: Involved in the conjugative transfer of the hemolysin
CC plasmid PAD1.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
KM Direct protein sequencing; Phenone.  
SQ SEQUENCE 8 AA; 819 MW; 04DD732C735B9C7 CRC64;  
Query Match 5.2%; Score 3; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.2e+06;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 33 FSL 35  
DB 2 FSL 4  
RESULT 12  
FAR1\_PANRE STANDARD; PRT; 8 AA.  
ID FAR1\_PANRE  
AC P41872;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE FMRamide-like neuropeptide PPI (SDPFLRF-amide).  
OS Panagrellus redivivus.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;  
OC Panagrolaimoidea; Panagrolaimidae; Panagrellus.  
OX NCBI\_TaxId=6233;  
RN [1]  
RP PROTEIN SEQUENCE.  
RA MEDLINE=33027659; PubMed=1408999; DOI=10.1016/0196-9781(92)90098-N;  
RA Geary T.G., Price D.A., Bowman J.W., Winterrowd C.A., Mackenzie C.D.,  
RA Garrison R.D., Williams J.F., Friedman A.R.;  
RT "Two FMRamide-like peptides from the free-living nematode Panagrellus  
RT redivivus";  
RL Peptides 13:209-214(1992).  
CC -1- FUNCTION: Myoactive.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- TISSUE SPECIFICITY: Nerve cords and paired groups of cells located  
CC caudally to the base of the pharynx.  
CC -1- SIMILARITY: Belongs to the FARP (FMRamide related peptide)  
CC family.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
KM Annotation: Direct protein sequencing; Neuropeptide.  
FT MOD RES 8 Phenylalanine amide.  
SQ SEQUENCE 8 AA; 995 MW; C6D40729C4576AB5 CRC64;  
Query Match 5.2%; Score 3; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.2e+06;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 29 FSL 31  
DB 5 FSL 7  
RESULT 13  
FAR1\_PENMO STANDARD; PRT; 8 AA.  
ID FAR1\_PENMO  
AC P83316;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE FMRamide-like neuropeptide FLPI (GPRNPLRF-amide).  
OS Penaeus monodon (Penaeid shrimp).  
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
OC Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;

OC Penaeidae; Penaeus.  
OX NCBI\_TaxId=6687;  
RN [1]  
RP PROTEIN SEQUENCE, AND MASS SPECTROMETRY.  
RC TISSUE=Eyestalk;  
RX MEDLINE=21956277; PubMed=11959015; DOI=10.1016/S1096-4959(01)00499-7;  
RA Sithigornkul P., Pupum J., Krungkarn C., Longyant S.,  
RA Chaivuthangkura P., Sithigornkul W., Petsom A.;  
RT "Seven novel FMRamide-like neuropeptide sequences from the eyestalk  
RT of the giant tiger prawn Penaeus monodon";  
RL Comp. Biochem. Physiol. 131B:325-337(2002).  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- MASS SPECTROMETRY: MW=1024.8; METHOD=MALDI; RANGE=1-8; NOTE=Ref.1.  
CC -1- SIMILARITY: Belongs to the FARP (FMRamide related peptide)  
CC family.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
DR GO:0007218; P:neuropeptide signaling pathway; TMS.  
KM Annotation: Direct protein sequencing; Neuropeptide.  
FT MOD RES 8 Phenylalanine amide.  
SQ SEQUENCE 8 AA; 1024 MW; 72D40729C4540AA8 CRC64;  
Query Match 5.2%; Score 3; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred. No. 2.2e+06;  
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 29 FSL 31  
DB 5 FSL 7  
RESULT 14  
FAR2\_MACRS STANDARD; PRT; 8 AA.  
ID FAR2\_MACRS  
AC P83275;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE FMRamide-like neuropeptide FLPI (ADKNPLRF-amide).  
OS Macrobrachium rosenbergii (Giant fresh water prawn).  
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Caridea;  
OC Palaemonoidea; Palaemonidae; Macrobrachium.  
OX NCBI\_TaxId=79674;  
RN [1]  
RP PROTEIN SEQUENCE, AND MASS SPECTROMETRY.  
RC TISSUE=Eyestalk;  
RA Sithigornkul P., Sarathongkum W., Jaidechaoey S., Longyant S.,  
RA Sithigornkul W.;  
RT "Novel FMRamide-like neuropeptides from the eyestalk of the giant  
RT freshwater prawn Macrobrachium rosenbergii";  
RL Comp. Biochem. Physiol. 120B:587-595(1998).  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- MASS SPECTROMETRY: MW=1009.4; METHOD=MALDI; RANGE=1-8; NOTE=Ref.1.  
CC -1- SIMILARITY: Belongs to the FARP (FMRamide related peptide)  
CC family.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
DR GO:0007218; P:neuropeptide signaling pathway; TMS.  
KM Annotation: Direct protein sequencing; Neuropeptide.  
FT MOD RES 8 Phenylalanine amide.  
SQ SEQUENCE 8 AA; 1010 MW; 9CD40729C4433AD CRC64;

Query Match 5.2%; Score 3; DB 1; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+06;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 29 FLR 31  
 |||  
 Db 5 FLR 7

## RESULT 15

FAR3\_HOMAM STANDARD; PRT; 8 AA.

ID FAR3\_HOMAM  
 AC P41486;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE FMRFamide-like neuropeptide 3 (Fli 3) (P2).  
 OS Homarus americanus (American lobster).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;  
 OC Nephropoidea; Nephropidae; Homarus.  
 OX NCBI\_TaxID=6706;  
 RP PROTEIN SEQUENCE.  
 RC TISSUE=Pericardial organs;  
 RX MEDLINE=8816164; PubMed=3429714;  
 RA Trimmer B.A., Koblitz L.A., Kravitz E.A.;  
 RT "Purification and characterization of FMRFamide-like immunoreactive  
 substances from the lobster nervous system: isolation and sequence  
 analysis of two closely related peptides.";  
 RT J. Comp. Neurol. 266:16-26(1987).  
 RL J. Comp. Neurol. 266:16-26(1987).  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- MISCELLANEOUS: Pericardial organs release this peptide with 100 mM  
 CC potassium in the presence of calcium.  
 CC -1- SIMILARITY: Belongs to the FMRP (FMRFamide related peptide)  
 CC family.

CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.

CC -----  
 CC Amidation; Direct protein sequencing; Neuropeptide.  
 KW MOD RES 8 Phenylalanine amide.  
 FT  
 SQ SEQUENCE 8 AA; 1054 MW; C6D40729C4540AB5 CRC64;

Query Match 5.2%; Score 3; DB 1; Length 8;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+06;  
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 29 FLR 31  
 |||  
 Db 5 FLR 7

Search completed: December 13, 2005, 08:25:36  
 Job time : 228 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2005, 08:18:37 ; Search time 46 Seconds  
(without alignments)  
104.243 Million cell updates/sec

Title: US-10-031-158B-14

Perfect score: 58  
Sequence: 1 MCMFPSPPLFFQLKQSS.....RYIGKRRATRFWDPRGTP 58

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 572060 seqs, 82675679 residues

Word size : 0

Total number of hits satisfying chosen parameters: 62763

Minimum DB seq length: 8  
Maximum DB seq length: 10

Post-processing: Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5 COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/6 COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/H.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/PCTUS.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/RB.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5	8.6	9	1	US-08-621-803-221
2	5	8.6	9	1	US-08-621-803-221
3	5	8.6	9	1	US-08-621-259A-213
4	5	8.6	9	2	US-08-159-339A-343
5	5	8.6	9	2	US-08-159-339A-405
6	5	8.6	9	2	US-09-217-352-221
7	5	8.6	9	2	US-09-677-664B-213
8	5	8.6	9	2	US-09-792-480-18
9	5	8.6	10	1	US-08-248-839C-135
10	5	8.6	10	2	US-09-648-400A-18
11	5	8.6	10	2	US-09-839-329-4
12	5	8.6	10	2	US-10-209-421-18
13	5	8.6	10	2	US-10-007-761-9
14	5	8.6	10	4	PCT-US96-01735-4
15	4	6.9	8	1	US-08-037-579A-6
16	4	6.9	8	1	US-08-244-855-6
17	4	6.9	8	1	US-08-529-190B-70
18	4	6.9	8	1	US-08-529-190B-71
19	4	6.9	8	1	US-08-922-267A-68
20	4	6.9	8	2	US-08-601-184-6
21	4	6.9	8	2	US-08-444-818-262
22	4	6.9	8	2	US-08-444-818-263
23	4	6.9	8	2	US-08-444-818-264
24	4	6.9	8	2	US-08-296-791-8
25	4	6.9	8	2	US-09-419-826-39
26	4	6.9	8	2	US-08-975-614-3
27	4	6.9	8	2	US-09-648-400A-10

28	4	6.9	8	2	US-08-475-955-194	Sequence 194, App
29	4	6.9	8	2	US-08-475-955-195	Sequence 195, App
30	4	6.9	8	2	US-08-475-955-196	Sequence 196, App
31	4	6.9	8	2	US-08-475-955-197	Sequence 197, App
32	4	6.9	8	2	US-08-475-955-204	Sequence 204, App
33	4	6.9	8	2	US-09-839-996-8	Sequence 8, Appl1
34	4	6.9	8	2	US-09-792-480-9	Sequence 9, Appl1
35	4	6.9	8	2	US-10-083-889-18	Sequence 18, Appl1
36	4	6.9	8	2	US-10-080-505-54	Sequence 54, Appl1
37	4	6.9	8	2	US-10-209-421-10	Sequence 10, Appl1
38	4	6.9	8	2	US-09-308-140-8	Sequence 8, Appl1
39	4	6.9	8	2	US-10-645-655-8	Sequence 8, Appl1
40	4	6.9	8	2	US-09-856-070-7	Sequence 7, Appl1
41	4	6.9	8	2	US-09-496-391-26	Sequence 26, Appl1
42	4	6.9	8	2	US-10-028-056-25	Sequence 25, Appl1
43	4	6.9	8	2	US-07-867-819D-139	Sequence 139, App
44	4	6.9	8	2	US-07-867-819D-141	Sequence 141, App
45	4	6.9	8	2	US-07-867-819D-142	Sequence 142, App

## ALIGNMENTS

RESULT 1  
US-08-621-803-221  
Sequence 221, Application US/08621803  
Patent No. 5851802  
GENERAL INFORMATION:  
APPLICANT: Belter, Marc D.  
TITLE OF INVENTION: Methods for Recombinant Microbial Production of  
TITLE OF INVENTION: Fusion Proteins and BPI-Derived Peptides  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/621,803  
FILING DATE: 22-MAR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Borun, Michael F.  
REGISTRATION NUMBER: 25,447  
REFERENCE/DOCKET NUMBER: 27129/33199  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ. ID NO.: 221:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: misc.feature  
OTHER INFORMATION: "XMP.382:  
NAME/KEY: Modified-site  
LOCATION: C-Terminus  
OTHER INFORMATION: /label= Amidation  
OTHER INFORMATION: /note= "The C-Terminus is Amidated."  
US-08-621-803-221  
Query Match 8.6%; Score 5; DB 1; Length 9;  
Best local Similarity 100.0%; Pred. No. 4.6e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 LQLLK 17  
|||||  
Db 4 LQLLK 8

RESULT 2  
US-08-621-259A-213  
Sequence 213, Application US/08621259A  
Patent No. 5858974  
GENERAL INFORMATION:  
APPLICANT: Little II, Roger G  
APPLICANT: Lim, Edward  
APPLICANT: Padem, Mitchell B.  
TITLE OF INVENTION: Anti-Fungal Peptides  
NUMBER OF SEQUENCES: 252  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrew, Held & Malloy, Ltd.  
STREET: 500 West Madison Street  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/621,259A  
FILING DATE: 21-MAR-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/504,841  
FILING DATE: 20-JUL-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 11021US02  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX:  
INFORMATION FOR SEQ ID NO: 213:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: "XMP.382:  
NAME/KEY: Modified-site  
LOCATION: C-terminus  
OTHER INFORMATION: //label= Amidation  
OTHER INFORMATION: /note= "The C-terminus is Amidated."  
US-08-621-259A-213

Query Match 8.6%; Score 5; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.6e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 LQLLK 17  
|||||  
Db 4 LQLLK 8

RESULT 3  
US-08-159-339A-343  
Sequence 343, Application US/08159339A  
Patent No. 6037135  
GENERAL INFORMATION:

APPLICANT: Kubo, Ralph T.  
APPLICANT: Grey, Howard M.  
APPLICANT: Sette, Alessandro  
APPLICANT: Celis, Esben  
TITLE OF INVENTION: HLA Binding peptides and Their  
TITLE OF INVENTION: Uses  
NUMBER OF SEQUENCES: 1254  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/159,339A  
FILING DATE: 29-NOV-1993  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/926,666  
FILING DATE: 07-AUG-1992  
APPLICATION NUMBER: US 08/027,746  
FILING DATE: 05-MAR-1993  
APPLICATION NUMBER: US 08/103,396  
FILING DATE: 06-AUG-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Ellen Lauer  
REGISTRATION NUMBER: 32,762  
REFERENCE/DOCKET NUMBER: 018623-005030US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
TELEX:  
INFORMATION FOR SEQ ID NO: 343:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-159-339A-343

Query Match 8.6%; Score 5; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.6e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 34 SLMLL 38  
|||||  
Db 1 SLMLL 5

RESULT 4  
US-08-159-339A-405  
Sequence 405, Application US/08159339A  
Patent No. 6037135  
GENERAL INFORMATION:  
APPLICANT: Kubo, Ralph T.  
APPLICANT: Grey, Howard M.  
APPLICANT: Sette, Alessandro  
APPLICANT: Celis, Esben  
TITLE OF INVENTION: HLA Binding peptides and Their  
TITLE OF INVENTION: Uses  
NUMBER OF SEQUENCES: 1254  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: CA

COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/159,339A  
FILING DATE: 29-NOV-1993  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/926,666  
FILING DATE: 07-AUG-1992  
APPLICATION NUMBER: US 08/027,746  
FILING DATE: 05-MAR-1993  
APPLICATION NUMBER: US 08/103,396  
FILING DATE: 06-AUG-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Ellen Lauver  
REGISTRATION NUMBER: 32,762  
REFERENCE/DOCKET NUMBER: 018623-005030US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
TELEX:  
INFORMATION FOR SEQ ID NO: 405:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-159-339A-405

Query Match 8.6%; Score 5; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.6e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 SLMML 38  
Db 3 SLMML 7

RESULT 5  
US-09-217-352-221  
Sequence 221, Application US/09217352  
Patent No. 6274344  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
TITLE OF INVENTION: Methods for Recombinant Microbial Production of  
TITLE OF INVENTION: Fusion Proteins and BPI-Derived Peptides  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/217,352  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/621,803  
FILING DATE: 22-MAR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Borun, Michael F.

REGISTRATION NUMBER: 25,447  
REFERENCE/DOCKET NUMBER: 27129/33199  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 221:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: "XMP.382:  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: C-Terminus  
OTHER INFORMATION: /label= Amidation  
OTHER INFORMATION: /note= "The C-Terminus is Amidated."  
US-09-217-352-221

Query Match 8.6%; Score 5; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.6e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LQLLK 17  
Db 4 LQLLK 8

RESULT 6  
US-09-677-664B-213  
Sequence 213, Application US/09677664B  
Patent No. 6664231  
GENERAL INFORMATION:  
APPLICANT: Little II, Roger G  
Liu, Edward  
Padem, Mitchell B.  
TITLE OF INVENTION: Anti-Fungal Peptides  
NUMBER OF SEQUENCES: 257  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/677,664B  
FILING DATE: 07-Mar-2003  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/227,659  
FILING DATE: 08-Jan-1999  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 11021US06  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/775-8000  
TELEFAX: 312/775-8100  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 213:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide

```

;
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: "XMP.382:"
;
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: C-Terminus
; OTHER INFORMATION: /label= Amidation
; /note= "The C-Terminus is Amidated."
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 213:
US-09-677-664B-213

Query Match      8.6%; Score 5; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      13 LQLLK 17
        |||||
Db      4 LQLLK 8

RESULT 7
US-09-792-480-18
; Sequence 18, Application US/09792480
; Patent No. 6669951
; GENERAL INFORMATION:
; APPLICANT: Rothbard, Jonathan B.
; APPLICANT: Wender, Paul A.
; APPLICANT: McGrane, P. Leo
; APPLICANT: Siesta, Lalitha V.S.
; APPLICANT: Kirschberg, Thorsten A.
; APPLICANT: Cellgate, Inc.
; TITLE OF INVENTION: Compositions and Methods for Enhancing Drug Delivery
; TITLE OF INVENTION: Across and Into Epithelial Tissues
; FILE REFERENCE: 019801-000230US
; CURRENT FILING DATE: US/09/792,480
; PRIOR APPLICATION NUMBER: US 09/648,400
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/150,510
; PRIOR FILING DATE: 1999-08-24
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:A-54
; OTHER INFORMATION: Ala-substituted analog of HIV-1 tat protein basic
; OTHER INFORMATION: region tat-49-57
; NAME/KEY: MOD_RES
; LOCATION: (1)
; OTHER INFORMATION: Xaa = fluorescein linked to amino group of
; OTHER INFORMATION: aminohexanoic acid (Fl-ahx) attached to the
; OTHER INFORMATION: N-terminal amino group of Arg
US-09-792-480-18

Query Match      8.6%; Score 5; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      43 KKRRA 47
        |||||
Db      2 KKRRA 6

RESULT 8
US-08-248-839C-135
; Sequence 135, Application US/08248839C
; Patent No. 5843702
; GENERAL INFORMATION:
; APPLICANT: McConnell, David
; APPLICANT: Devine, Kevin
```

```

;
; APPLICANT: O'Kane, Charles
; TITLE OF INVENTION: A Gene Expression System
; NUMBER OF SEQUENCES: 185
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 58437020 No. 5843702disk of No. 5843702th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/248,839C
; FILING DATE: 25-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gregg, Valeta A.
; REGISTRATION NUMBER: 35,127
; REFERENCE/DOCKET NUMBER: 3614.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 135:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
US-08-248-839C-135

Query Match      8.6%; Score 5; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      34 SIMLL 38
        |||||
Db      3 SIMLL 7

RESULT 9
US-09-648-400A-18
; Sequence 18, Application US/09648400A
; Patent No. 6593292
; GENERAL INFORMATION:
; APPLICANT: Rothbard, Jonathan B.
; APPLICANT: Wender, Paul A.
; APPLICANT: McGrane, P. Leo
; APPLICANT: Siesta, Lalitha V.S.
; APPLICANT: Kirschberg, Thorsten A.
; APPLICANT: Cellgate, Inc.
; TITLE OF INVENTION: Compositions and Methods for Enhancing Drug Delivery
; TITLE OF INVENTION: Across and Into Epithelial Tissues
; FILE REFERENCE: 019801-000210US
; CURRENT FILING DATE: US/09/648,400A
; PRIOR APPLICATION NUMBER: US 60/150,510
; PRIOR FILING DATE: 1999-08-24
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:A-54
; OTHER INFORMATION: Ala-substituted analog of HIV-1 tat protein basic
; OTHER INFORMATION: region tat-49-57
; FEATURE:
```

NAME/KEY: MOD\_RES  
LOCATION: (1)  
OTHER INFORMATION: Xaa = fluorescein linked to amino group of  
OTHER INFORMATION: aminohexanoic acid (Pl-ahx)  
US-09-648-400A-18

Query Match  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 GKRR 47  
Db 3 KKRR 7

RESULT 10  
US-09-839-329-4  
Sequence 4, Application US/09839329  
Patent No. 6664040  
GENERAL INFORMATION:  
APPLICANT: Michael P. Sherman  
APPLICANT: Warner C. Greene  
APPLICANT: Carlos M.C. de No. 6664040hna  
APPLICANT: Ulrich Schubert  
APPLICANT: Peter Henklein  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DELIVERY OF  
FILE REFERENCE: G&C 30448.91-US-U2  
CURRENT APPLICATION NUMBER: US/09/839,329  
PRIOR FILING DATE: 2001-04-20  
PRIOR APPLICATION NUMBER: 60/206,610  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/267,827  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Human immunodeficiency virus  
US-09-839-329-4

Query Match  
Best Local Similarity 8.6%; Score 5; DB 2; Length 10;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 GKRR 46  
Db 2 GKRR 6

RESULT 11  
US-10-209-421-18  
Sequence 18, Application US/10209421  
Patent No. 6759387  
GENERAL INFORMATION:  
APPLICANT: Rothbard, Jonathan B.  
APPLICANT: Wender, Paul A.  
APPLICANT: McGee, P. Leo  
APPLICANT: Sista, Lalitha V.S.  
APPLICANT: Kirschberg, Thorsten A.  
APPLICANT: Celigae, Inc.  
TITLE OF INVENTION: Compositions and Methods for Enhancing Drug Delivery  
FILE REFERENCE: 019801-000211US  
CURRENT APPLICATION NUMBER: US/10/209,421  
PRIOR FILING DATE: 2002-07-30  
PRIOR APPLICATION NUMBER: US 60/150,510  
PRIOR FILING DATE: 1999-08-24  
PRIOR APPLICATION NUMBER: US 09/648,400  
PRIOR FILING DATE: 2000-08-24

NUMBER OF SEQ ID NOS: 51  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 18  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Artificial Sequence

FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence:A-54  
OTHER INFORMATION: Ala-substituted analog of HIV-1 tat protein basic  
OTHER INFORMATION: region Tat-49-57  
FEATURE:  
NAME/KEY: MOD\_RES  
LOCATION: (1)  
OTHER INFORMATION: Xaa = fluorescein linked to amino group of  
OTHER INFORMATION: aminohexanoic acid (Pl-ahx)  
US-10-209-421-18

Query Match  
Best Local Similarity 8.6%; Score 5; DB 2; Length 10;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 KKRR 47  
Db 3 KKRR 7

RESULT 12  
US-10-007-761-9  
Sequence 9, Application US/10007761  
Patent No. 6855693  
GENERAL INFORMATION:  
APPLICANT: Mochly-Rosen, Daria  
TITLE OF INVENTION: Peptides for Activation and Inhibition  
FILE REFERENCE: 58600-8208-US00  
CURRENT APPLICATION NUMBER: US/10/007,761  
CURRENT FILING DATE: 2001-11-09  
PRIOR APPLICATION NUMBER: US 60/262,060  
PRIOR FILING DATE: 2001-01-18  
NUMBER OF SEQ ID NOS: 72  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 9  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Tat-derived carrier peptide  
US-10-007-761-9

Query Match  
Best Local Similarity 8.6%; Score 5; DB 2; Length 10;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 GKRR 46  
Db 2 GKRR 6

RESULT 13  
PCT-US96-01735-4  
Sequence 4, Application PC/TUS9601735  
GENERAL INFORMATION:  
APPLICANT: Marks, Andrew R.  
TITLE OF INVENTION: HUMAN T CELL INOSITOL 1,4,5,-TRISPHOSPHATE  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10112-0228

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/01735  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/386,039  
FILING DATE: 09-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: KOLE, Lisa B  
REGISTRATION NUMBER: 35,225  
REFERENCE/DOCKET NUMBER: A30042 - 165/30555  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-408-2628  
TELEFAX: 212-765-2519  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
FEATURE:  
PCT-US96-01735-4  
Query Match 8.6%; Score 5; DB 4; Length 10;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 12 FLOLL 16  
Db 4 FLOLL 8  
RESULT 14  
US-08-037-579A-6  
Sequence 6, Application US/08037579A  
Patent No. 5552537  
GENERAL INFORMATION:  
APPLICANT: Zhang, Ke  
APPLICANT: Max, Edward E  
APPLICANT: Saxon, Andrew  
TITLE OF INVENTION: IGE ISOFORMS AND METHODS OF USE  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSER: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/037,579A  
FILING DATE: 24-MAR-1993  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Rowland, Berttram I  
REGISTRATION NUMBER: 20,015  
REFERENCE/DOCKET NUMBER: A-57950/BIR UCLA-233

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1589  
TELEFAX: (415) 398-3249  
TELEX: 910 277229 FHT UR  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-037-579A-6  
Query Match 6.9%; Score 4; DB 1; Length 8;  
Best Local Similarity 100.0%; Pred. No. 4.6e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 19 SSRR 22  
Db 2 SSRR 5  
RESULT 15  
US-08-244-855-6  
Sequence 6, Application US/08244855  
Patent No. 5762906  
GENERAL INFORMATION:  
APPLICANT: Creighton, Andrew M.  
TITLE OF INVENTION: FURTHER IMPROVEMENTS RELATING TO  
TITLE OF INVENTION: RADIO LABELLING OF PROTEINS  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSER: NIXON & VANDERHIE  
STREET: 8TH FLOOR, 1100 NORTH GLEBE ROAD  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/244,855  
FILING DATE: 16-JUN-1994  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Mitchard, Leonard C.  
REGISTRATION NUMBER: 29,009  
REFERENCE/DOCKET NUMBER: 604-285  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4005  
TELEFAX: (703) 816-4100  
TELEX: 200797 NIXN UR  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-244-855-6  
Qy 45 RRAT 48  
Db 2 RRAT 5  
Search completed: December 13, 2005, 08:26:27

Wed Dec 14 08:46:43 2005

us-10-031-158b-14.oliszm.ra1

Page 7

Job time : 47 secs

---

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2005, 08:25:42 ; Search time 159 Seconds  
(Without alignments)  
152.416 Million cell updates/sec

Title: US-10-031-158b-14

Perfect score: 58  
Sequence: 1 MGFPPSPPLFFFLQLKQSS.....RYGKKRRATRFWDPRGRTP 58

Scoring table: OLIGO  
Gapop 60.0, Gapext 60.0

Searched: 1867569 seqs, 417829326 residues

Word size: 0

Total number of hits satisfying chosen parameters: 151674

Minimum DB seq length: 8  
Maximum DB seq length: 10

Post-processing: Listing first 45 summaries

Database:

Published Applications AA Main:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5	8.6	8	US-10-137-867-292	Sequence 292, App
2	5	8.6	9	US-09-765-527-221	Sequence 221, App
3	5	8.6	9	US-09-792-480-18	Sequence 18, App1
4	5	8.6	9	US-09-870-216C-3	Sequence 3, App1
5	5	8.6	9	US-10-017-327-3	Sequence 3, App1
6	5	8.6	9	US-10-144-188-2	Sequence 2, App1
7	5	8.6	10	US-09-839-329-4	Sequence 4, App1
8	5	8.6	10	US-09-572-404B-238	Sequence 238, App
9	5	8.6	10	US-09-572-404B-1444	Sequence 1444, App
10	5	8.6	10	US-09-572-404B-3212	Sequence 3212, App
11	5	8.6	10	US-10-007-761-9	Sequence 9, App1
12	5	8.6	10	US-10-083-960-18	Sequence 18, App1
13	5	8.6	10	US-10-209-421-18	Sequence 18, App1
14	5	8.6	10	US-10-168-789A-26	Sequence 26, App1
15	5	8.6	10	US-10-168-789A-27	Sequence 27, App1
16	5	8.6	10	US-10-421-548-7	Sequence 7, App1
17	5	8.6	10	US-10-421-503-65	Sequence 65, App1
18	5	8.6	10	US-10-843-731-9	Sequence 9, App1
19	5	8.6	10	US-10-936-237-47	Sequence 47, App1
20	5	8.6	10	US-10-513-003-65	Sequence 65, App1
21	4	6.9	8	US-09-815-837-114	Sequence 114, App
22	4	6.9	8	US-09-804-682-170	Sequence 170, App
23	4	6.9	8	US-09-792-480-9	Sequence 9, App1
24	4	6.9	8	US-09-756-875-10	Sequence 10, App1
25	4	6.9	8	US-09-756-875-11	Sequence 11, App1
26	4	6.9	8	US-09-756-875-20	Sequence 20, App1
27	4	6.9	8	US-09-756-875-21	Sequence 21, App1

28	4	6.9	8	US-09-756-875-22	Sequence 22, App1
29	4	6.9	8	US-09-756-875-23	Sequence 23, App1
30	4	6.9	8	US-09-756-875-24	Sequence 24, App1
31	4	6.9	8	US-09-756-875-29	Sequence 29, App1
32	4	6.9	8	US-09-017-743C-7	Sequence 7, App1
33	4	6.9	8	US-09-943-123-19	Sequence 19, App1
34	4	6.9	8	US-09-839-996-8	Sequence 8, App1
35	4	6.9	8	US-09-726-470A-76	Sequence 76, App1
36	4	6.9	8	US-09-726-470A-137	Sequence 137, App
37	4	6.9	8	US-09-726-470A-244	Sequence 244, App
38	4	6.9	8	US-09-792-861A-15	Sequence 15, App1
39	4	6.9	8	US-09-770-875-15	Sequence 15, App1
40	4	6.9	8	US-09-876-904A-353	Sequence 353, App
41	4	6.9	8	US-09-876-904A-408	Sequence 408, App
42	4	6.9	8	US-09-876-904A-417	Sequence 417, App
43	4	6.9	8	US-09-876-904A-419	Sequence 419, App
44	4	6.9	8	US-09-796-076-10	Sequence 10, App1
45	4	6.9	8	US-10-028-056-25	Sequence 25, App1

#### ALIGNMENTS

RESULT 1  
US-10-137-867-292  
Sequence 292, Application US/10137867  
Publication No. US20030207349A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary B.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P330R1C146  
CURRENT APPLICATION NUMBER: US/10/137,867  
CURRENT FILING DATE: 2002-05-03  
Prior Application removed - See Palm or File Wrapper  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 292  
LENGTH: 428  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-137-867-292  
Query Match 8.6%; Score 5; DB 4; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Cv 9 LFPL 13  
Db 1 LFPL 5  
RESULT 2  
US-09-765-527-221  
Sequence 221, Application US/09765527  
Patent No. US2002000638A1  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.

TITLE OF INVENTION: Methods for Recombinant Microbial Production of Fusion Proteins and BPI-Derived Peptides  
NUMBER OF SEQUENCES: 265  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Marshall, O'Toole, Gerestein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/765,527  
FILING DATE: 18-Jan-2001  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/621,803  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Borun, Michael F.  
REGISTRATION NUMBER: 25,447  
REFERENCE/DOCKET NUMBER: 27129/33199  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 221:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: misc\_feature  
OTHER INFORMATION: "XMP.382:  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: C-Terminus  
OTHER INFORMATION: //label= Amidation  
/note= "The C-Terminus is Amidated."  
SEQUENCE DESCRIPTION: SEQ ID NO: 221:  
US-09-765-527-221  
Query Match 8.6%; Score 5; DB 3; Length 9;  
Best Local Similarity 100.0%; Pred.No. 1.7e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CY 13 LQLLK 17  
DB 4 LQLLK 8  
RESULT 3  
US-09-792-480-18  
Sequence 18, Application US/09792480  
Patent No. US20020127198A1  
GENERAL INFORMATION:  
APPLICANT: Rothbard, Jonathan B.  
APPLICANT: Wender, Paul A.  
APPLICANT: McGrane, P. Leo  
APPLICANT: Siera, Lalitha V.S.  
APPLICANT: Kirschberg, Thorsten A.  
TITLE OF INVENTION: Compositions and Methods for Enhancing Drug Delivery  
TITLE OF INVENTION: Across and Into Epithelial Tissues  
FILE REFERENCE: 019801-000230US  
CURRENT APPLICATION NUMBER: US/09/792,480  
CURRENT FILING DATE: 2001-02-23  
PRIOR APPLICATION NUMBER: US 09/648,400  
PRIOR FILING DATE: 2000-08-24

PRIOR APPLICATION NUMBER: US 60/150,510  
PRIOR FILING DATE: 1999-08-24  
NUMBER OF SEQ ID NOS: 57  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 18  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence:A-54  
OTHER INFORMATION: Ala-substituted analog of HIV-1 tat protein basic  
OTHER INFORMATION: region Tat-49-57  
NAME/KEY: MOD\_RES  
LOCATION: (1)  
OTHER INFORMATION: Xaa = fluorescein linked to amino group of  
OTHER INFORMATION: aminohexanoic acid (PI-ahx) attached to the  
OTHER INFORMATION: N-terminal amino group of Arg  
US-09-792-480-18  
Query Match 8.6%; Score 5; DB 3; Length 9;  
Best Local Similarity 100.0%; Pred.No. 1.7e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CY 43 KKRA 47  
DB 2 KKRA 6  
RESULT 4  
US-09-870-216C-3  
Sequence 3, Application US/09870216C  
Publication No. US20040138135A1  
GENERAL INFORMATION:  
APPLICANT: Charles A. Nicolette  
TITLE OF INVENTION: THERAPEUTIC COMPOUNDS FOR OVARIAN CANCER  
FILE REFERENCE: 68126881210100  
CURRENT APPLICATION NUMBER: US/09/870,216C  
PRIOR FILING DATE: 2001-05-30  
PRIOR APPLICATION NUMBER: 60/209,391  
PRIOR FILING DATE: 2000-05-31  
PRIOR APPLICATION NUMBER: 60/226,256  
PRIOR FILING DATE: 2000-08-17  
PRIOR APPLICATION NUMBER: 60/257,008  
PRIOR FILING DATE: 2000-12-20  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-870-216C-3  
Query Match 8.6%; Score 5; DB 3; Length 9;  
Best Local Similarity 100.0%; Pred.No. 1.7e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CY 12 FLQLL 16  
DB 1 FLQLL 5  
RESULT 5  
US-10-017-327-3  
Sequence 3, Application US/10017327  
Publication No. US20020155471A1  
GENERAL INFORMATION:  
APPLICANT: Charles A. Nicolette  
TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS AND  
TITLE OF INVENTION: METHODS FOR USING SAME  
FILE REFERENCE: GZ 2101.20  
CURRENT APPLICATION NUMBER: US/10/017,327  
CURRENT FILING DATE: 2001-12-06  
NUMBER OF SEQ ID NOS: 11

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-327-3
```

```
Query Match      8.6%; Score 5; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      12 FLOQL 16
        |||||
Db       1 FLOQL 5
```

```
RESULT 6
US-10-144-188-2
; Sequence 2, Application US/10144188
; Publication No. US20030170212A1
; GENERAL INFORMATION:
; APPLICANT: Cal, Zeling
; APPLICANT: Jackson, Michael R.
; APPLICANT: Peterson, Per A.
; APPLICANT: Shi, Weixing
; APPLICANT: Kong, Yan
; APPLICANT: Degraw, Juli
; TITLE OF INVENTION: Ex-Vivo Priming For Generating Cytotoxic T Lymphocytes Specific
; TITLE OF INVENTION: For No. US20030170212A1-Tumor Antigens To Treat Autoimmune And A
; FILE REFERENCE: PRI0010 ORT-1627
; CURRENT APPLICATION NUMBER: US/10/144,188
; CURRENT FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: 60/291,300
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide antigen
US-10-144-188-2
```

```
Query Match      8.6%; Score 5; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      5 PPSPL 9
        |||||
Db       2 PPSPL 6
```

```
RESULT 7
US-09-839-329-4
; Sequence 4, Application US/09839329
; Publication No. US20020022027A1
; GENERAL INFORMATION:
; APPLICANT: Michael P. Sherman
; APPLICANT: Warner C. Greene
; APPLICANT: Carlos M.C. de No. 66640400hna
; APPLICANT: Ulrich Schubert
; APPLICANT: Peter Henklein
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DELIVERY OF
; TITLE OF INVENTION: A MOLECULE INTO A CELL
; FILE REFERENCE: G&C 30448-91-US-U2
; CURRENT APPLICATION NUMBER: US/09/839,329
; CURRENT FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/206,610
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/267,827
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 5
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Human immunodeficiency virus
US-09-839-329-4
```

```
Query Match      8.6%; Score 5; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      42 GKRR 46
        |||||
Db       2 GKRR 6
```

```
RESULT 8
US-09-572-404B-238
; Sequence 238, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 238
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in MTNR1B at 351-360 and may interact with Sequer
US-09-572-404B-238
```

```
Query Match      8.6%; Score 5; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      20 SRRLE 24
        |||||
Db       1 SRRLE 5
```

```
RESULT 9
US-09-572-404B-1444
; Sequence 1444, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 1444
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in NTAK at 681-690 and may interact with Sequence
; OTHER INFORMATION: in this patent.
US-09-572-404B-1444
```

```
Query Match      8.6%; Score 5; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 53 PRGRT 57  
Db 6 PRGRT 10

RESULT 10  
US-09-572-404B-3212  
; Sequence 3212, Application US/09572404B  
; Publication No. US20030078374A1  
; GENERAL INFORMATION:  
; APPLICANT: Proteom Ltd  
; TITLE OF INVENTION: Complementary peptide ligands from the human genome  
; FILE REFERENCE: Human patent  
; CURRENT APPLICATION NUMBER: US/09/572,404B  
; CURRENT FILING DATE: 2000-05-17  
; NUMBER OF SEQ ID NOS: 4203  
; SOFTWARE: ProPatent version 1.0  
; SEQ ID NO 3212  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; OTHER INFORMATION: sequence located in HOXA9 OR HOXA10 at 369-378 and may interact w/1  
US-09-572-404B-3212

Query Match 8.6%; Score 5; DB 3; Length 10;  
Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 RRATR 49  
Db 4 RRATR 8

RESULT 11  
US-10-007-761-9  
; Sequence 9, Application US/10007761  
; Publication No. US20020150984A1  
; GENERAL INFORMATION:  
; APPLICANT: Mochly-Rosen, Darla  
; TITLE OF INVENTION: Peptides for Activation and Inhibition  
; FILE REFERENCE: Of delta-PKC  
; FILE REFERENCE: 58600-8208.US00  
; CURRENT APPLICATION NUMBER: US/10/007,761  
; CURRENT FILING DATE: 2001-11-09  
; PRIOR APPLICATION NUMBER: US 60/262,060  
; PRIOR FILING DATE: 2001-01-18  
; NUMBER OF SEQ ID NOS: 72  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Tat-derived carrier peptide  
US-10-007-761-9

Query Match 8.6%; Score 5; DB 4; Length 10;  
Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 42 GKRR 46  
Db 2 GKRR 6

RESULT 12  
US-10-083-960-18  
; Sequence 18, Application US/10083960  
; Publication No. US20030022831A1  
; GENERAL INFORMATION:  
; APPLICANT: Rothbard, Jonathan B.

APPLICANT: Wender, Paul A.  
; APPLICANT: McGrane, P. Leo  
; APPLICANT: Sista, Lalitha V.S.  
; APPLICANT: Kirschberg, Thorsten A.  
; APPLICANT: CellGate, Inc.  
; TITLE OF INVENTION: Compositions and Methods for Enhancing  
; TITLE OF INVENTION: Drug Delivery Across and Into Ocular Tissues  
; FILE REFERENCE: 019801-000240US  
; CURRENT APPLICATION NUMBER: US/10/083,960  
; CURRENT FILING DATE: 2003-07-14  
; PRIOR APPLICATION NUMBER: US 60/150,510  
; PRIOR FILING DATE: 1999-08-24  
; PRIOR APPLICATION NUMBER: US 09/648,400  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 09/792,480  
; PRIOR FILING DATE: 2001-02-23  
; NUMBER OF SEQ ID NOS: 86  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 18  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: A-54 alanine-substituted analog of Tat-49-57  
; NAME/KEY: MOD\_RES  
; LOCATION: (1)...(1)  
; OTHER INFORMATION: Xaa = fluorescein conjugated aminohexanoic acid  
; OTHER INFORMATION: (Fl-ahx)  
US-10-083-960-18

Query Match 8.6%; Score 5; DB 4; Length 10;  
Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 KKRA 47  
Db 3 KKRA 7

RESULT 13  
US-10-209-421-18  
; Sequence 18, Application US/10209421  
; Publication No. US20030083256A1  
; GENERAL INFORMATION:  
; APPLICANT: Rothbard, Jonathan B.  
; APPLICANT: Wender, Paul A.  
; APPLICANT: McGrane, P. Leo  
; APPLICANT: Sista, Lalitha V.S.  
; APPLICANT: Kirschberg, Thorsten A.  
; APPLICANT: CellGate, Inc.  
; TITLE OF INVENTION: Compositions and Methods for Enhancing Drug Delivery  
; TITLE OF INVENTION: Across and Into Epithelial Tissues  
; FILE REFERENCE: 019801-000211US  
; CURRENT APPLICATION NUMBER: US/10/209,421  
; CURRENT FILING DATE: 2002-07-30  
; PRIOR APPLICATION NUMBER: US 60/150,510  
; PRIOR FILING DATE: 1999-08-24  
; PRIOR APPLICATION NUMBER: US 09/648,400  
; PRIOR FILING DATE: 2000-08-24  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:A-54  
; OTHER INFORMATION: Ala-substituted analog of HIV-1 tat protein basic  
; OTHER INFORMATION: region Tat-49-57  
; NAME/KEY: MOD\_RES  
; LOCATION: (1)

; OTHER INFORMATION: Xaa = fluorescein linked to amino group of  
; OTHER INFORMATION: aminohexanoic acid (Fl-ahx)  
US-10-209-421-18

Query Match 8.6%; Score 5; DB 4; Length 10;  
Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 KKRRA 47  
Db 3 KKRRA 7

RESULT 14  
US-10-168-789A-26

; Sequence 26, Application US/10168789A  
; Publication No. US20030148943A1  
; GENERAL INFORMATION:  
; APPLICANT: ITOH, Yasuaki  
; APPLICANT: NISHI, Kazuo  
; APPLICANT: KITADA, Chieko  
; APPLICANT: INATONI, No. US20030148943A1uhiro  
; TITLE OF INVENTION: No. US20030148943A1el Tachykinin-like Polypeptides and Use Thereof  
; FILE REFERENCE: 2680USOP  
; CURRENT APPLICATION NUMBER: US/10/168, 789A  
; PRIOR FILING DATE: 2002-06-20  
; PRIOR APPLICATION NUMBER: PCT/JP00/09083  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: JP 11-362638  
; PRIOR FILING DATE: 1999-12-21  
; PRIOR APPLICATION NUMBER: JP 12-066714  
; PRIOR FILING DATE: 1999-03-10  
; NUMBER OF SEQ ID NOS: 64  
; SEQ ID NO 26  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: peptide  
US-10-168-789A-26

Query Match 8.6%; Score 5; DB 4; Length 10;  
Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 LEHTF 27  
Db 2 LEHTF 6

RESULT 15  
US-10-168-789A-27

; Sequence 27, Application US/10168789A  
; Publication No. US20030148943A1  
; GENERAL INFORMATION:  
; APPLICANT: ITOH, Yasuaki  
; APPLICANT: NISHI, Kazuo  
; APPLICANT: KITADA, Chieko  
; APPLICANT: INATONI, No. US20030148943A1uhiro  
; TITLE OF INVENTION: No. US20030148943A1el Tachykinin-like Polypeptides and Use Thereof  
; FILE REFERENCE: 2680USOP  
; CURRENT APPLICATION NUMBER: US/10/168, 789A  
; PRIOR FILING DATE: 2002-06-20  
; PRIOR APPLICATION NUMBER: PCT/JP00/09083  
; PRIOR FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: JP 11-362638  
; PRIOR FILING DATE: 1999-12-21  
; PRIOR APPLICATION NUMBER: JP 12-066714  
; PRIOR FILING DATE: 1999-03-10  
; NUMBER OF SEQ ID NOS: 64  
; SEQ ID NO 27  
; LENGTH: 10  
; TYPE: PRT

; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Artificial Polypeptide  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (01)..(01)  
; OTHER INFORMATION: Xaa means pyroglutamic acid  
US-10-168-789A-27

Query Match 8.6%; Score 5; DB 4; Length 10;  
Best Local Similarity 100.0%; Pred. No. 7.4e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 LEHTF 27  
Db 2 LEHTF 6

Search completed: December 13, 2005, 08:39:30  
Job time : 160 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 13, 2005, 08:26:32 ; Search time 11 Seconds  
(without alignments)  
29.445 Million cell updates/sec

Title: US-10-031-158B-14

Perfect score: 58  
Sequence: 1 MGFPPSPLEFFILQKQSS.....RYIGKKRRATRFWDPRGRGP 58

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 32527 seqs, 5584426 residues

Word size: 0

Total number of hits satisfying chosen parameters: 3786

Minimum DB seq length: 8  
Maximum DB seq length: 10

Post-processing: Listing first 45 summaries

Database:

1: Published Applications AA New:  
2: /cgn2\_6/prodata/2/pubppa/US09\_NEW\_PUB.pep:  
3: /cgn2\_6/prodata/2/pubppa/US07\_NEW\_PUB.pep:  
4: /cgn2\_6/prodata/2/pubppa/US08\_NEW\_PUB.pep:  
5: /cgn2\_6/prodata/2/pubppa/PC7\_NEW\_PUB.pep:  
6: /cgn2\_6/prodata/2/pubppa/US10\_NEW\_PUB.pep:  
7: /cgn2\_6/prodata/2/pubppa/US11\_NEW\_PUB.pep:  
8: /cgn2\_6/prodata/2/pubppa/US60\_NEW\_PUB.pep:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5	8.6	10	US-11-119-098-8	Sequence 8, Appl1
2	4	6.9	8	US-10-914-842A-26	Sequence 26, Appl1
3	4	6.9	9	US-10-499-715-15	Sequence 15, Appl1
4	4	6.9	9	US-10-499-715-16	Sequence 16, Appl1
5	4	6.9	9	US-11-176-868-1	Sequence 1, Appl1
6	4	6.9	9	US-11-097-864-144	Sequence 144, App
7	4	6.9	9	US-11-097-912-144	Sequence 144, App
8	4	6.9	10	US-10-444-662-1	Sequence 1, Appl1
9	4	6.9	10	US-11-152-747-28	Sequence 28, Appl1
10	4	6.9	10	US-11-032-794-53	Sequence 53, Appl1
11	4	6.9	10	US-11-097-864-446	Sequence 446, App
12	4	6.9	10	US-11-097-864-511	Sequence 511, App
13	4	6.9	10	US-11-097-912-446	Sequence 446, App
14	4	6.9	10	US-11-097-912-511	Sequence 511, App
15	4	6.9	10	US-11-102-432-1	Sequence 1, Appl1
16	3	5.2	8	US-10-997-066-22	Sequence 22, Appl1
17	3	5.2	8	US-10-839-966-6	Sequence 6, Appl1
18	3	5.2	8	US-10-467-657-7372	Sequence 7372, Ap
19	3	5.2	8	US-10-467-657-8729	Sequence 8729, Ap
20	3	5.2	8	US-10-467-657-8769	Sequence 8769, Ap
21	3	5.2	8	US-10-467-657-8797	Sequence 8797, Ap
22	3	5.2	8	US-10-914-842A-17	Sequence 17, Appl1
23	3	5.2	8	US-10-416-047-5	Sequence 5, Appl1
24	3	5.2	8	US-11-082-251-14	Sequence 14, Appl1
25	3	5.2	8	US-11-096-706-208	Sequence 208, App

#### ALIGNMENTS

26	3	5.2	8	US-11-176-868-17	Sequence 17, Appl1
27	3	5.2	8	US-11-032-498-24	Sequence 24, Appl1
28	3	5.2	8	US-11-152-747-32	Sequence 32, Appl1
29	3	5.2	8	US-11-054-515-3141	Sequence 3141, Ap
30	3	5.2	8	US-11-054-515-3214	Sequence 3214, Ap
31	3	5.2	8	US-11-058-727-98	Sequence 98, Appl1
32	3	5.2	8	US-11-058-735-8	Sequence 8, Appl1
33	3	5.2	8	US-11-058-735-9	Sequence 9, Appl1
34	3	5.2	8	US-11-058-735-10	Sequence 10, Appl1
35	3	5.2	8	US-11-108-389-98	Sequence 98, Appl1
36	3	5.2	8	US-11-101-287-152	Sequence 152, Appl1
37	3	5.2	8	US-11-115-922-213	Sequence 213, App
38	3	5.2	9	US-10-997-066-8	Sequence 8, Appl1
39	3	5.2	9	US-10-997-066-10	Sequence 10, Appl1
40	3	5.2	9	US-10-499-715-17	Sequence 17, Appl1
41	3	5.2	9	US-10-499-715-18	Sequence 18, Appl1
42	3	5.2	9	US-10-981-873-26	Sequence 26, Appl1
43	3	5.2	9	US-10-073-301A-10	Sequence 10, Appl1
44	3	5.2	9	US-10-952-535A-44	Sequence 44, Appl1
45	3	5.2	9	US-10-467-657-9002	Sequence 9002, Ap

```
RESULT 1
US-11-119-098-8
; Sequence 8, Application US/1119098
; Publication No. US20050267030A1
; GENERAL INFORMATION:
; APPLICANT: Taao, Philip S.
; TITLE OF INVENTION: Use of deltaPKC Peptides for Modulation of Reactive Oxygen Species
; FILE REFERENCE: 58600-8213-US00
; CURRENT APPLICATION NUMBER: US/11/119, 098
; PRIOR FILING DATE: 2005-04-29
; PRIOR APPLICATION NUMBER: US 60/567, 315
; PRIOR FILING DATE: 2004-04-30
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Tat-derived carrier peptide
US-11-119-098-8

Query Match      8.6%  Score 5; DB 7; Length 10;
Best Local Similarity 100.0%; Pred.No. 7.9;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      42 GKRR 46
Db      2 GKRR 6

RESULT 2
US-10-914-842A-26
; Sequence 26, Application US/10914842A
; Publication No. US20050260730A1
; GENERAL INFORMATION:
; APPLICANT: FISCHER, PETER MARTIN
; TITLE OF INVENTION: CDK2/CYCLIN A CRYSTALS AND USES THEREOF
; FILE REFERENCE: CCI-032
; CURRENT APPLICATION NUMBER: US/10/914, 842A
; PRIOR FILING DATE: 2004-08-10
; PRIOR APPLICATION NUMBER: UK 0324465.4
; PRIOR FILING DATE: 2003-10-20
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 26
; LENGTH: 8
```

```

; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-10-914-842A-26
```

```

Query Match
Best Local Similarity 100.0%; Score 4; DB 6; Length 8;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 44 KRRR 47
Db 3 KRRR 6
```

```

RESULT 3
US-10-499-715-15
; Sequence 15, Application US/10499715
; Publication No. US20050250717A1
; GENERAL INFORMATION:
; APPLICANT: BENAVAHU, Dafna
; APPLICANT: SHUR, Iritna
; TITLE OF INVENTION: CHROMATIN REMODELING PROTEIN AS A MARKER EXPRESSED BY STROMAL PRO
; FILE REFERENCE: BENAVAHU-1.1 PCT
; CURRENT APPLICATION NUMBER: US/10/499,715
; CURRENT FILING DATE: 2004-06-21
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 15
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-499-715-15
```

```

Query Match
Best Local Similarity 100.0%; Score 4; DB 6; Length 9;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 43 KRRR 46
Db 1 KRRR 4
```

```

RESULT 4
US-10-499-715-16
; Sequence 16, Application US/10499715
; Publication No. US20050250717A1
; GENERAL INFORMATION:
; APPLICANT: BENAVAHU, Dafna
; APPLICANT: SHUR, Iritna
; TITLE OF INVENTION: CHROMATIN REMODELING PROTEIN AS A MARKER EXPRESSED BY STROMAL PRO
; FILE REFERENCE: BENAVAHU-1.1 PCT
; CURRENT APPLICATION NUMBER: US/10/499,715
; CURRENT FILING DATE: 2004-06-21
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 16
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-499-715-16
```

```

Query Match
Best Local Similarity 100.0%; Score 4; DB 6; Length 9;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 43 KRRR 46
Db 1 KRRR 4
```

```

RESULT 5
US-11-176-868-1
; Sequence 1, Application US/1176868
; Publication No. US20050245454A1
; GENERAL INFORMATION:
; APPLICANT: Goldstein, Gideon
; TITLE OF INVENTION: Methods and Compositions for Impairing Multiplication of HIV-1
; FILE REFERENCE: GGP3USA
; CURRENT APPLICATION NUMBER: US/11/176,868
; CURRENT FILING DATE: 2005-07-08
; PRIOR APPLICATION NUMBER: US/10/323,013
; PRIOR FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: US/10/114,176
; PRIOR FILING DATE: 2002-04-02
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-11-176-868-1
```

```

Query Match
Best Local Similarity 100.0%; Score 4; DB 7; Length 9;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 43 KRRR 46
Db 2 KRRR 5
```

```

RESULT 6
US-11-097-864-144
; Sequence 144, Application US/11097864
; Publication No. US20050265924A1
; GENERAL INFORMATION:
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Faries, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN NAMED 161P2F10B
; FILE REFERENCE: USEFUL IN THE TREATMENT AND DETECTION OF CANCER
; CURRENT APPLICATION NUMBER: US/11/097,864
; CURRENT FILING DATE: 2005-04-01
; PRIOR APPLICATION NUMBER: US 10/062,109
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 10/005,480
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 765
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 144
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo Sapien
US-11-097-864-144
```

```

Query Match
Best Local Similarity 100.0%; Score 4; DB 7; Length 9;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 32 NFSL 35
Db 6 NFSL 9
```

```

RESULT 7
US-11-097-912-144
; Sequence 144, Application US/11097912
; Publication No. US20050265921A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Challita-Bid, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Paris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 161P2F10B
; TITLE OF INVENTION: USEFUL IN TREATMENT AND DETECTION OF CANCER
; FILE REFERENCE: 511582006204
; CURRENT APPLICATION NUMBER: US/11/097,912
; CURRENT FILING DATE: 2005-04-01
; PRIOR APPLICATION NUMBER: US 10/062,109
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 10/005,480
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 765
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 144
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo Sapien
US-11-097-912-144
```

```
Query Match          6.9%; Score 4; DB 7; Length 10;
Best Local Similarity 100.0%; Pred. No. 2.8e+04;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy          32 NFSL 35
Db          6 NFSL 9
```

```
RESULT 8
US-10-444-662-1
; Sequence 1, Application US/10444662
; Publication No. US20050250683A9
; GENERAL INFORMATION:
; APPLICANT: Mirus Corporation
; APPLICANT: Rozema, David
; APPLICANT: Wolff, Jon
; APPLICANT: Wakefield, Darren
; APPLICANT: Ekens, Kirk
; APPLICANT: Hestrom, James
; TITLE OF INVENTION: Reversible Modification of Membrane Interaction
; FILE REFERENCE: Mirus.035.01
; CURRENT APPLICATION NUMBER: US/10/444,662
; CURRENT FILING DATE: 2003-05-23
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-10-444-662-1
```

```
Query Match          6.9%; Score 4; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy          43 KKRR 46
Db          3 KKRR 6
```

```
RESULT 9
US-11-152-747-28
; Sequence 28, Application US/11152747
; Publication No. US20050251881A1
; GENERAL INFORMATION:
; APPLICANT: E. I. du Pont de Nemours, Inc.
; APPLICANT: Cheng, Qiong
; APPLICANT: Tao, Luan
```

```
; TITLE OF INVENTION: CAROTENOID KETOYLASE GENE
; FILE REFERENCE: CL-1849 US NA
; CURRENT APPLICATION NUMBER: US/11/152,747
; CURRENT FILING DATE: 2005-06-14
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 28
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Rhodococcus erythropolis AN12
US-11-152-747-28
```

```
Query Match          6.9%; Score 4; DB 7; Length 10;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy          38 LRYI 41
Db          3 LRYI 6
```

```
RESULT 10
US-11-032-794-53
; Sequence 53, Application US/11032794
; Publication No. US20050265970A1
; GENERAL INFORMATION:
; APPLICANT: FRANZOSO, GUIDO
; APPLICANT: PAPA, SALVATORE
; APPLICANT: BUBICI, CONCETTA
; APPLICANT: DESMAELE, ENRICO
; APPLICANT: ZAZZERONI, FRANCESCA
; TITLE OF INVENTION: IDENTIFICATION OF NOVEL FACTORS THAT BLOCK PROGRAMMED
; FILE REFERENCE: 21459-97816
; CURRENT APPLICATION NUMBER: US/11/032,794
; CURRENT FILING DATE: 2005-01-10
; PRIOR APPLICATION NUMBER: 11/000,365
; PRIOR FILING DATE: 2004-11-29
; PRIOR APPLICATION NUMBER: 60/526,231
; PRIOR FILING DATE: 2003-12-02
; PRIOR APPLICATION NUMBER: 10/626,905
; PRIOR FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: 10/263,330
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 60/328,911
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/326,492
; PRIOR FILING DATE: 2001-10-02
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 53
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 1
US-11-032-794-53
```

```
Query Match          6.9%; Score 4; DB 7; Length 10;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy          43 KKRR 46
Db          3 KKRR 6
```

```
RESULT 11
US-11-097-864-446
; Sequence 446, Application US/11097864
; Publication No. US20050265924A1
; GENERAL INFORMATION:
; APPLICANT: Challita-Bid, Pia M.
```

```

: APPLICANT: Raitano, Arthur B.
: APPLICANT: Paris, Mary
: APPLICANT: Hubert, Rene S.
: APPLICANT: Morrison, Karen Jane Meyrick
: APPLICANT: Jakobovits, Aya
: TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN NAMED 161P2F10B
: TITLE OF INVENTION: USEFUL IN THE TREATMENT AND DETECTION OF CANCER
: FILE REFERENCE: 511582006205
: CURRENT APPLICATION NUMBER: US/11/097,864
: CURRENT FILING DATE: 2005-04-01
: PRIOR APPLICATION NUMBER: US 10/062,109
: PRIOR FILING DATE: 2002-01-31
: PRIOR APPLICATION NUMBER: US 10/005,480
: PRIOR FILING DATE: 2001-11-07
: NUMBER OF SEQ ID NOS: 765
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 446
: LENGTH: 10
: TYPE: PRT
: ORGANISM: Homo Sapien
US-11-097-864-446
```

```

Query Match          6.9%; Score 4; DB 7; Length 10;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy      32 NFSL 35
        ||||
Db      7 NFSL 10
```

```

RESULT 12
US-11-097-864-511
: Sequence 511, Application US/11097864
: Publication No. US20050265924A1
: GENERAL INFORMATION:
: APPLICANT: Chailita-Bid, Pia M.
: APPLICANT: Raitano, Arthur B.
: APPLICANT: Paris, Mary
: APPLICANT: Hubert, Rene S.
: APPLICANT: Morrison, Karen Jane Meyrick
: APPLICANT: Jakobovits, Aya
: TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN NAMED 161P2F10B
: TITLE OF INVENTION: USEFUL IN THE TREATMENT AND DETECTION OF CANCER
: FILE REFERENCE: 511582006205
: CURRENT APPLICATION NUMBER: US/11/097,864
: CURRENT FILING DATE: 2005-04-01
: PRIOR APPLICATION NUMBER: US 10/062,109
: PRIOR FILING DATE: 2002-01-31
: PRIOR APPLICATION NUMBER: US 10/005,480
: PRIOR FILING DATE: 2001-11-07
: NUMBER OF SEQ ID NOS: 765
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 511
: LENGTH: 10
: TYPE: PRT
: ORGANISM: Homo Sapien
US-11-097-864-511
```

```

Query Match          6.9%; Score 4; DB 7; Length 10;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy      32 NFSL 35
        ||||
Db      7 NFSL 10
```

```

RESULT 13
US-11-097-912-446
: Sequence 446, Application US/11097912
: Publication No. US20050265921A1
: GENERAL INFORMATION:
```

```

: APPLICANT: Chailita-Bid, Pia M.
: APPLICANT: Raitano, Arthur B.
: APPLICANT: Paris, Mary
: APPLICANT: Hubert, Rene S.
: APPLICANT: Morrison, Karen Jane Meyrick
: APPLICANT: Jakobovits, Aya
: TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 161P2F10B
: TITLE OF INVENTION: USEFUL IN TREATMENT AND DETECTION OF CANCER
: FILE REFERENCE: 511582006204
: CURRENT APPLICATION NUMBER: US/11/097,912
: CURRENT FILING DATE: 2005-04-01
: PRIOR APPLICATION NUMBER: US 10/062,109
: PRIOR FILING DATE: 2002-01-31
: PRIOR APPLICATION NUMBER: US 10/005,480
: PRIOR FILING DATE: 2001-11-07
: NUMBER OF SEQ ID NOS: 765
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 446
: LENGTH: 10
: TYPE: PRT
: ORGANISM: Homo Sapien
US-11-097-912-446
```

```

Query Match          6.9%; Score 4; DB 7; Length 10;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy      32 NFSL 35
        ||||
Db      7 NFSL 10
```

```

RESULT 14
US-11-097-912-511
: Sequence 511, Application US/11097912
: Publication No. US20050265921A1
: GENERAL INFORMATION:
: APPLICANT: Chailita-Bid, Pia M.
: APPLICANT: Raitano, Arthur B.
: APPLICANT: Paris, Mary
: APPLICANT: Hubert, Rene S.
: APPLICANT: Morrison, Karen Jane Meyrick
: APPLICANT: Jakobovits, Aya
: TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN ENTITLED 161P2F10B
: TITLE OF INVENTION: USEFUL IN TREATMENT AND DETECTION OF CANCER
: FILE REFERENCE: 511582006204
: CURRENT APPLICATION NUMBER: US/11/097,912
: CURRENT FILING DATE: 2005-04-01
: PRIOR APPLICATION NUMBER: US 10/062,109
: PRIOR FILING DATE: 2002-01-31
: PRIOR APPLICATION NUMBER: US 10/005,480
: PRIOR FILING DATE: 2001-11-07
: NUMBER OF SEQ ID NOS: 765
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 511
: LENGTH: 10
: TYPE: PRT
: ORGANISM: Homo Sapien
US-11-097-912-511
```

```

Query Match          6.9%; Score 4; DB 7; Length 10;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy      32 NFSL 35
        ||||
Db      7 NFSL 10
```

```

RESULT 15
US-11-102-432-1
: Sequence 1, Application US/11102432
: Publication No. US20050265957A1
```

; GENERAL INFORMATION:  
; APPLICANT: Mirus Bio  
; APPLICANT: Monahan, Sean  
; APPLICANT: Nader, Lisa  
; TITLE OF INVENTION: Polymerized Formamides for Use in Delivery of Compounds to Cells  
; FILE REFERENCE: Mirus.044.01  
; CURRENT APPLICATION NUMBER: US//11/102,432  
; CURRENT FILING DATE: 2005-04-08  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Human immunodeficiency virus  
US-11-102-432-1

Query Match 6.9%; Score 4; DB 7; Length 10;  
Best Local Similarity 100.0%; Pred. No. 89;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 KKR 46  
|||  
Db 3 KKR 6

Search completed: December 13, 2005, 08:39:48  
Job time : 12 secs

**THIS PAGE BLANK (USPTO)**